

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 12:19:58 ; Search time 7822 Seconds  
(without alignments)  
8945.197 Million cell updates/sec

Title: US-09-868-744B-1  
Perfect score: 1444  
Sequence: 1 attatgatctcaatacatt.....gtactggaatgtaggatcc 1444

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1442	99.9	1444	6	AX026532
2	1349.6	93.5	75508	8	AB026654 Arabidops
3	1341.8	92.9	6385	6	AX463618 Sequence
4	1341.4	92.9	4526	6	AX463617 Sequence
5	1306.2	90.5	11127	6	AX392814 Sequence
6	1300.8	90.1	14184	6	CQ774496 Sequence
7	1298.4	89.9	3450	6	AX463620 Sequence
8	1298.4	89.9	7474	6	CQ814616 Sequence
9	1298.4	89.9	9356	6	CQ814630 Sequence
10	1297.4	89.8	3408	6	AX463619 Sequence
11	1297	89.8	2857	6	AX463621 Sequence
12	1234.2	85.5	3172	8	ATU41998
13	1184	82.0	1217	6	AX461227 Sequence
14	1174.8	81.4	1219	6	AR236191 Sequence
15	1174.8	81.4	1219	6	AR438153 Sequence
16	1174.8	81.4	1219	6	AX167634 Sequence
17	1174	81.3	1742	6	AR236212 Sequence
18	1174	81.3	1742	6	AR438174 Sequence
19	1174	81.3	1742	6	AX167655 Sequence

20	1078.8	74.7	15676	6	BD251968	BD251968	Artificia
21	1078.8	74.7	17111	6	BD251969	BD251969	Artificia
22	1078.8	74.7	17116	6	BD251970	BD251970	Artificia
23	707.2	49.0	8573	6	CQ821161	CQ821161	Sequence
24	475.8	33.0	11461	6	BD266633	BD266633	Regulatio
25	475.8	33.0	11461	6	AX047314	AX047314	Sequence
26	474.8	32.9	12766	6	BD266630	BD266630	Regulatio
27	474.8	32.9	12766	6	AX047311	AX047311	Sequence
28	163	11.3	1486	8	AY087751	AY087751	Arabidops
29	158	10.9	1679	8	AY064043	AY064043	Arabidops
30	158	10.9	2637	8	AY056800	AY056800	Arabidops
31	153	10.6	1559	8	AF428330	AF428330	Arabidops
c	32	96.8	6.7	119091	8	AC016041	Genomic s
c	33	95	6.6	122	8	ATH524576	Arabidops
34	94.4	6.5	3261	8	ATU42007	U42007	Arabidopsis
35	76	5.3	2000	6	AX508607	AX508607	Sequence
36	75.4	5.2	1271	6	AR236192	AR236192	Sequence
37	75.4	5.2	1271	6	AR438154	AR438154	Sequence
38	75.4	5.2	1271	6	AX167635	AX167635	Sequence
39	73.8	5.1	1800	6	AR236211	AR236211	Sequence
40	73.8	5.1	1800	6	AR438173	AR438173	Sequence
41	73.8	5.1	1800	6	AX167654	AX167654	Sequence
42	69.8	4.8	255	11	CR376614	CR376614	Arabidops
43	62.6	4.3	2460	8	STPOAC71	X55750	S. tuberosum
44	62.2	4.3	253001	3	AE014834	AE014834	Plasmodium
c	45	61.6	4.3	251448	3	AE014819	Plasmodium

ALIGNMENTS

RESULT 1  
AX026532  
LOCUS AX026532 1444 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 1 from Patent WO0037661.  
ACCESSION AX026532  
VERSION AX026532.1 GI:10187719  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
REFERENCE 1  
AUTHORS Pertj, J.H., Schepers, F.M. and van Dun Cornelis, M.P.  
TITLE Genetic modification of compositae  
JOURNAL Patent: WO 0037661-A 1 29-JUN-2000;  
ADVANTA TECHNOLOGY LIMITED (GB) ; PERTIJS JAN H (NL) ; DUN CORNELIS M P VAN (NL) ; SCHEPERS FRANK M A (NL)  
FEATURES  
source 1. 1444  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN	Query Match 99.9%; Score 1442; DB 6; Length 1444; Best Local Similarity 100.0%; Pred. No. 5.1e-301; Matches 1444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATTATGATCTCAATATACATATATCTCATCTAGATCTAGGTATCATTTATGTAAAG 60 
Db	1 ATTATGATCTCAATATACATATCTCATCTAGATCTAGGTATCATTTATGTAAAG 60 
Qy	61 AAAGTTTTCAGCAATATGNNACGACAAATGGCTACCTCGATGTAATTTGGTATCTCAAC 120 
Db	61 AAAGTTTTCAGCAATATGNNACGACAAATGGCTACCTCGATGTAATTTGGTATCTCAAC 120 
Qy	121 TCAACATTATCTTATACCAACATTAGTTAGCAAAATTTAAACACTATTTTATGTAT 180 
Db	121 TCAACATTATCTTATACCAACATTAGTTAGCAAAATTTAAACACTATTTTATGTAT 180 
Qy	181 GCAAGAGTCAGCATATGTATAATTGATTTCAGAAATCGTTTTTCGACGATTCGGATGTAGTAG 240 

181 GCAAGAGTCAGCATATGTATTAATTTGATTGAGAAATCGTTTGGACGAGTTCCGGATGTAGTAG 240  
241 TAGCCATTATTTAAATGTACATACTAATCGTGAATAGTATGATGATGAACATCTGATCTT 300  
241 TAGCCATTATTTAAATGTACATACTAATCGTGAATAGTATGATGATGAACATCTGATCTT 300  
301 ATTGTATATAATATCCATAAACAACATCATGAAGAAGACATTTCTTCCAGGCTCTGAATTAAT 360  
301 ATTGTATATAATATCCATAAACAACATCATGAAGAAGACATTTCTTCCAGGCTCTGAATTAAT 360  
361 TATGATACAAATCTTAATAGAAAAACGAATTAATTAAGTTGAATTTGATGAATCTAAATTG 420  
361 TATGATACAAATCTTAATAGAAAAACGAATTAATTAAGTTGAATTTGATGAATCTAAATTG 420  
421 AACAGCCAAACACGACGAGGACTAATCGTTGCTGGATGATGACTCGGTTTAAGTTAACAC 480  
421 AACAGCCAAACACGACGAGGACTAATCGTTGCTGGATGATGACTCGGTTTAAGTTAACAC 480  
481 TAAAAAAACGGAGCTGTCATGTAAACACGCGGATCGAGCAGGTACAGTCAATGAAGCCATC 540  
481 TAAAAAAACGGAGCTGTCATGTAAACACGCGGATCGAGCAGGTACAGTCAATGAAGCCATC 540  
541 AAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATTAAGTTTAAAAATTAAGTTAAAC 600  
541 AAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATTAAGTTTAAAAATTAAGTTAAAC 600  
601 GAGGAAAGCTGTCGACAGCAGGTACGTTATCTTTACCTGTGTGTCGAAATGAATCG 660  
601 GAGGAAAGCTGTCGACAGCAGGTACGTTATCTTTACCTGTGTGTCGAAATGAATCG 660  
661 TGTCTGTCCATTTTAATTAATTTTGAAGCCGAAATAAAGTTGTAAGAGATAAACC 720  
661 TGTCTGTCCATTTTAATTAATTTTGAAGCCGAAATAAAGTTGTAAGAGATAAACC 720  
721 CGCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGCTCGTTGCTCCTCTCAC 780  
721 CGCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGCTCGTTGCTCCTCTCAC 780  
781 TTTTCATCAGCGTTTGAATCTCCGCGACTTGACAGAGAAGAACAGAGAAGACTAA 840  
781 TTTTCATCAGCGTTTGAATCTCCGCGACTTGACAGAGAAGAACAGAGAAGACTAA 840  
841 GAGAGAACTAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTTCCTCAATCTC 900  
841 GAGAGAACTAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTTCCTCAATCTC 900  
901 ATCTTCTTCTCCGCTCTTTTCTTCCAAAGGTAATAGGAATCTTCTGGATCTACTTTATTT 960  
901 ATCTTCTTCTCCGCTCTTTTCTTCCAAAGGTAATAGGAATCTTCTGGATCTACTTTATTT 960  
961 GCTGGATCTGATCTGTTTCTCAATTTCTTGAATCTGGAATCTGGAATCTGTTTAAATTGGAT 1020  
961 GCTGGATCTGATCTGTTTCTCAATTTCTTGAATCTGGAATCTGGAATCTGTTTAAATTGGAT 1020  
1021 CTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGATCTAGTTGACCGATCAGT 1080  
1021 CTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGATCTAGTTGACCGATCAGT 1080  
1081 TAGCTCGATTAGCTACCAAGATTTGGCTTGACCTTGATGGAGAGATCCATGTTTCATGT 1140  
1081 TAGCTCGATTAGCTACCAAGATTTGGCTTGACCTTGATGGAGAGATCCATGTTTCATGT 1140  
1141 TACCTGGGAAATGATTTGTATATGTGAATCTGAATCTGAATCTGTTGAAGTTAGATTGAAT 1200  
1141 TACCTGGGAAATGATTTGTATATGTGAATCTGAATCTGAATCTGTTGAAGTTAGATTGAAT 1200  
1201 CTGAACACTGTAAGTTAGATTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTGT 1260  
1201 CTGAACACTGTAAGTTAGATTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTGT 1260  
1261 ATAGATTCTTCCAAACCTTAGGATTTGTAGTGTGTCGATCGTTGAACAGAAAGCTATTTCTG 1320

Db 1261 ATAGATTCTTCCAAACCTTAGGATTTGTAGTGTGTCGATCGTTGAACAGAAAGCTATTTCTG 1320  
Qy 1321 ATTCAAATCAGGGTTTATTTGACTGTATTGAATCTCTTTTGTGTGTTTCAGCTCATATAAA 1380  
Db 1321 ATTCAAATCAGGGTTTATTTGACTGTATTGAATCTCTTTTGTGTGTTTCAGCTCATATAAA 1380  
Qy 1381 AATGGCTGAGGCTGACGATATTCAACCAATCGTGTGACCAATGGTACTGGAATCGTAGG 1440  
Db 1381 AATGGCTGAGGCTGACGATATTCAACCAATCGTGTGACCAATGGTACTGGAATCGTAGG 1440  
Qy 1441 ATCC 1444  
Db 1441 ATCC 1444  
RESULT 2  
AB026654  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S.  
Structural analysis of Arabidopsis thaliana chromosome 3. I.  
Sequence features of the regions of 4,504,864 bp covered by sixty  
P1 and TAC clones  
DNA Res. 7 (2), 131-135 (2000)  
20277480  
10819329  
2 (bases 1 to 75508)  
Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S.  
Direct Submission  
Submitted (28-APR-1999) Yasukazu Nakamura, Kazusa DNA Research  
Institute, Department of Plant Gene Research; 1532-3, Yana,  
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,  
Tel: 81-438-52-3935, Fax: 81-438-52-3934)  
Address for correspondence: kaos@kazusa.or.jp  
For the latest information on annotation of this clone, please see  
http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=MVE11  
Genes with similarity to proteins in the databases are described in  
'product' or 'note' qualifiers. Genes that have no significant  
protein similarity are described as 'unknown protein'.  
The software programs used to predict genes include: Grail  
(Informatics Group, Oak Ridge National Laboratory,  
http://compbio.ornl.gov/Grail-1.3/),  
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),  
NetGene2 (S.M. Hebsgaard, et al., CBS, technical University of  
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
SplicePredictor (Volker Brendel, Stanford University,  
http://grem1inl.2001.iastate.edu/cgi-bin/sp.cgi).  
Genes encoding tRNAs are predicted by tRNAscan-SE  
(Sean Eddy, Washington University School of Medicine, St. Louis,  
http://genome.wustl.edu/eddy/tRNAscan-SE/).  
This sequence may not be the entire insert of this clone. It may be  
shorter because we remove overlaps between neighboring submissions.  
The 5' clone is K24M9 and the 3' clone is MCB22.  
Location/Qualifiers  
1..75508  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="MVE11"  
/clone\_lib="Mitsui P1"  
/ecotype="Columbia"  
/note="unnamed protein product; gene\_id:MVE11.1"  
FEATURES  
source  
CDS

unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01791.1"  
/db\_xref="GI:9293888"  
/translation="MASSSSSSLSFSTSKKNTFKKNSAFSSSRATSLIKRQOTVP  
KAKELSLICDDIVCVICVSGNGELKTWEEREKVATARRIGELSETKRRGKSVDLH  
EFLKRNKDDPEKEBKIKVRRVPKVKYFVMDPRFDYVSEQLMQLVQSLERNLRI  
QRTCAVBAQOORVQVYMANQBLMANTWNLQOHSNQSYMVLWHNGGAFSQIP  
VSALASNTQSLAPIPPELMIYPNSDAGNYSGLGVQGTGINGLQNMNMLTYNNINSV  
NDFSQFQONSRAEYSYSLGVHEDGNNEFENPNMNNFNVDQACGLLHGQAGTN  
GLQSMWHDYSNNINSGLSHQYVQPTYNQHQDRVFNLDONGNTRSL"  
complement(join(3045..3317,3774..4546,4652..5271,  
5392..5628,5717..5793))  
/note="unnamed protein product; gb|AAC34345.1  
gene\_id:MVE11.2  
strong similarity to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01792.1"  
/db\_xref="GI:9293889"  
/translation="MANGPAAPATTTTGGDSRRRLSASIEAICRRFRNRKSGGGRS  
DMVKPFNINFTQDKNSCCCTFKFOIVKLLFILLSATLFTIISPEAVHSHS  
SRWLRQDPYRFDLDINDDDVTKLENIIEGRTIIVLNFDSNEIORWEVRSKSD  
NGDEKVVNLNDYADKNVMDALYPEWIDEEQTEVPVCPNIPNIKRRLDLIV  
KLPCKEGNSRDVGRLLQLAAATVAASAKGFFRGVFFVSRCPPIPNLFRCKOLVS  
RRGDVLYKNLDTLRLDLQPLVSGCELSPLGIQDRPSLGNPKREAYATILHSAHV  
YCAAAQSRQSGSTRDLVLVDNITSGVHRSGLAAGWOIRTIQRNPKAEKDA  
VNEGNSFRLWOLYDKIPIFIDALLILRNIDFLSMPELSATGNGTLPNSGMV  
LEPCNTQLLMEHNEIESYNGGQGYLNEVFTWHRIIPKHMFLKHFVIGDEDDAK  
IKTSLFGAEPVLYVLYHGLMKWLCYRDYDCNFNSDIFVEFATDIAHRKNNWVHDA  
MPQELQCYLRSKQAQLEVDYRQAAEANYADGHWKIRVKDPREKICIDKLCNNKSM  
LRHWGESNWDYVESVPTPAITVDRSSSLPGHNL"  
complement(join(7969..8583,8684..9123,9207..9332,  
9437..9586,9675..10305))  
/note="unnamed protein product; emb|CAB70981.1  
gene\_id:MVE11.3  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01793.1"  
/db\_xref="GI:9293890"  
/translation="MDTEKGYPRGFRASLRPETGPRFRFDSMOIDBAADLQKDEIR  
QENSYLVFNKIDSGELEATQFLDRNPEALTALTITNGDPTPHKAVLSGHIKIVE  
IIRIHDEQVLYKINDNGYALTAYAAAGVIRIAECLVKNCPGIVSRNAKEHIPIV  
VARSYGHVLYVLYSHTPSLDPCDDSDHKGNGAMLVNCTIVDGLYCTALDLIQ  
RYPKLAYTRSDNDTAIMALACTPYAFPSGTFLAFQWRIYSCIHIEKINNPHEVNH  
HHHQYRNSQBNSSQEBLKHLYLFRIRIRVYKLGHAQAKEILDICQELPKF  
DAAQQNAGLQALPQAVENGIVEYIEEMRHYPDIVWSKNSSGNIIFYAVSORQEK  
IFSLYNTGAKKNILATNWDIFHNMLHHAAYRAPASRLNIPGAALQMQBELQWPK  
VEKLQPKHRKVNVLKQKTPKALFDQHKDQVQEGEKWKETATCTVVAALITWM  
FSAFTVPGYRSDCMPLYIHOHRPKIFLISDAISLFTSCMSLLMFLGILKSRYEED  
FLRSPTKLIVGLLALFLSMATMIVFVVTWLTVLGEXISVWSAQFMFLAVIPGMFV  
VLQFPVLEIIPATYCPNVFKPRVFKL"  
join(11367..11581,11799..11940,12035..12088,12453..12544,  
12656..12882,12966..13142,13242..13354)  
/note="unnamed protein product; contains similarity to  
uridylylate kinase  
gene\_id:MVE11.4"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01794.1"  
/db\_xref="GI:9293891"  
/translation="MAIPLTSCSPISTSSISRTSFVPLTRNRTFTFSQWYRRV  
LISCSSLSGSSGSDPSMNGNGSSLGQSPFLPSPDGTSKPPLKQWRVLLK  
VSGEALADGEONIDPKVTALAREVAATRLGIEVAIVVGGNIFRGSTWAGCSGLD  
RSSADYIGMLATVMAIFLQATMESIGIPTRVQTAFRMSVAPYIRRAIRHLEKGR  
WVIFAAGTNPFTTDTAALRCAEINAEVVLKATNVDFVDDPFRPNRNLHLSLT  
YQEVTSKOLSVMDMTAILTCQENNIPIVVVFNLSGPCNTAKAIKGRVCTGLIGTWSI  
VTTTS"  
complement(13589..14257)  
/note="unnamed protein product; gb|AAF16598.1  
gene\_id:MVE11.5

similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01795.1"  
/db\_xref="GI:9293892"  
/translation="MDPSEYFAGGNPSDOONKROQLQICGPRPSPLSVHKDSHKIKKP  
PKHPAPPNRQPPYIPREPVIIVAYSPKVHATASEFMNVVQRLTGISGVSFLSG  
GGDVSPPAARLASTENASPGKEPAARDETVEINTAMEAAEFGYAGFGLSPSPAL  
LPDTAGTGFSPMYHQGMFSPAIPGLGSPAGFMSPPRSPGFTSLVASPTFADFSTH  
WDQST"  
complement(join(15556..15731,16089..16154,16246..16333,  
16599..16739))  
/note="unnamed protein product; gene\_id:MVE11.6  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01796.1"  
/db\_xref="GI:9293893"  
/translation="MPTSYTSPSAOPTNNWFLFNIFSSLLVIALRFFDDQDVSDRY  
SYKONGLFLNENAFVKKQIGLKFNKIITNMKBKAVLVENEIGLKFNKIKFVKGBEY  
PCPSLPALPDTKASSNSDLDPAPNHEYQVSVNSPTLLNDSTGEVSKGPCWSHLL"  
complement(18068..19115)  
/note="unnamed protein product; gene\_id:MVE11.7  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01797.1"  
/db\_xref="GI:9293894"  
/translation="MGRDETETITVPFFKCPISLDVMSRSPVSLCTGYTVDRASIQR  
WIDGQNTCPATMQLLKTQDFVNLTLQRLINISDSIGRRHNGSPLVNPSPGREVP  
TKBEYNLLERLSLENLMKIVFVKDSDSNREFLSKMBFVPMVLVDIIRTKTKIEL  
VIMAIRIDISKVDRERLNLMLANDGGDCLTALLAIQORGLKIESVRVLDWISF  
DAKSLMIAERDGLVTENMKISITESSDPSLIEASLSFLITISKSRRVSKLIAAKA  
ITKILILLTETNVAVTEKSLKLETLTSSKREGLRICGDNRCRVEGVVKLLKV  
STTAHAVTILMCLCYVEREDKTVETVERSNGVTKLLVLVIQSNCSAMVRQAKDLI  
KVLNPSALLAAEYKTHIMPFF"  
complement(join(28267..29317,29494..29585))  
/note="unnamed protein product; gb|AAD43172.1  
gene\_id:MVE11.8  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01798.1"  
/db\_xref="GI:9293895"  
/translation="WTRRTQYPIAESITARSIVQALPASATISGNGPKKKKNCVN  
RGLMDKOIPTLLOELISRLGKANIHASLVCKTWLKEAVSRKQSPMLPYPQSOR  
GGPEGDYLVNPSRQTHLKFELTYGRNKLACAKDGLLVKXNDPVDVFLNPFPT  
GERICLPQPNSTRDCLTFSAAPTSTSCCVISFTPQSFYAVVKDVTWRPGESVMT  
HFQPKRYGEVINCIFSGNMFYCLSTSGRLSVDFPDSRETWNLPVKPCRAFRKIML  
VRQVMEHEGDI FVVTTRVNNRKLAFKLQNVNWEEMKVPNGLTVFSSDATSLT  
RAGLPEERNLISDIDDFVKSSHPTFYDYDCSAWLQPHDNFNF"  
complement(join(29896..30389,30465..30582,30745..30879,  
31009..31129,31210..31276,31364..31468,31677..31788,  
31869..31981,32106..32223,32374..32474,32694..32810,  
33023..33152,33271..34000,34112..34210,34701..34792,  
35000..35176,35688..36488,36613..36879))  
/note="unnamed protein product; gene\_id:MVE11.9  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01799.1"  
/db\_xref="GI:9293896"  
/translation="MGRDVAARAKYRKABEVGRDRRQARWANNVDILKNHGYVD  
ALKWFRIDYDLSVKYLPCKDILLPTCOSIGETVLRLENPEEALIVYKHLQLAERANDT  
VEKORACTQLGRTTHEMPLKEDCEALQSAKKYFKKAMELAQILKEKPPGESSGFL  
BEYNANNNIGMLDLDNPEARTILKKGLQICDEEVRDYDAARSRLHNLGNVFM  
ALRSDWAKAGHIENDIINCHIKNVQGEAKYQINLAHLNKTQYIDALLCYGKASSL  
AKSMQDSALVEQIENHTKIVKSMKMEELREBELMLKCLSAEMTDAGTSEERKSM  
LOVNAKLSGLDKSSVMVFLQYSKRKKISDELCDKEKLSGAPMIVGESYQNLK  
NFRSLKWFIRSTYEGHEAIGNLEQALAKINIGLDCIGETWGTALQVBEGRYALK  
ANLPSIQSALIEDIHYHMRFGNAQKASELKTIONLKESEHAKEKASCSTQDECS

Best Local Similarity 98.7%; Pred. No. 2.8e-281; Matches 1423; Conservative 0; Mismatches 11; Indels 8; Gaps 6;																		
QY	1	ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATATGTAAAG	60															
Db	57986	ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATATGTAAAG	58045															
QY	61	AAAGTTTTCACGAATATGNNACGACAAAATGGCTACACTCGATGTAAATGGGTATCTCAAC	120															
Db	58046	AAAGTTTTCACGAATATGNCACGACAAAATGGCTAGACTCGATGTAAATGGGTATCTCAAC	58105															
QY	121	TCACATTTATCTATPACCAACATTTAGTTAG-CAAAATTTAAACAACCTA-TTTTATATGT	178															
Db	58106	TCACATTTATCTATPACCAACATTTAGTTAGCAAAAATTTAAACAACCTATTTTTATGT	58165															
QY	179	ATGCAAGAGTCAGCATATGTATTAATTTGATTCGAATCGTTTTGACGAGTTCGGATGTAGT	238															
Db	58166	ATGCAAGAGTCAGCATATGTATTAATTTGATTCGAATCGTTTTGACGAGTTCGGATGTAGT	58225															
QY	239	AGTAGCCATTATTTAATGTACATCTAAATCGTGAATAGTG-ATATGATCAAAACATTTGTAT	297															
Db	58226	AGTAGCCATTATTTAATGTACATCTAAATCGTGAATAGTGATGATGAAACATTTGTAT	58285															
QY	298	CTTATTGTATAAATPCCATAACACATCATGAAAGACACTTTCTTTCAAGGCTCTGAATT	357															
Db	58286	CTTATTGTATAAATPCCATAAACACATCATGAAAGACACTTTCTTTCAAGGCTCTGAATT	58345															
QY	358	AATTATGATCAAAATCTTAATAGAAAACGAATTAATTTAGTTTGAATTTGTATGAAATCTTAA	417															
Db	58346	AATTATGATCAAAATCTTAATAGAAAACGAATTAATTTAGTTTGAATTTGTATGAAATCTTAA	58405															
QY	418	TTGAACAAGCCAACACGACGAGGACTAAGCTTTGCTGATTTGACTCGGTTTAAAGTTTAAAC	477															
Db	58406	TTGAACAAGCCAACACGACGAGGACTAAGCTTTGCTGATTTGACTCGGTTTAAAGTTTAAAC	58465															
QY	478	CACATAAAAAACGGAGCTGTATGTAAACACGCGGATCGAGCAGGTACAGTCTATGAAGCC	537															
Db	58466	CACATAAAAAACGGAGCTGTATGTAAACACGCGGATCGAGCAGGTACAGTCTATGAAGCC	58525															
QY	538	ATCAAAAGCAAAAGAACTAATCCNAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	597															
Db	58526	ATCAAAAGCAAAAGAACTAATCCNAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	58585															
QY	598	CACAGGGGAAAA-GCTGTCTGACAGCCAGGTACAGTTATCTTTTACCTGTGGTCGAAATCA	656															
Db	58586	CACAGGGGAAAAAGGCTGTCTGACAGCCAGGTACAGTTATCTTTTACCTGTGGTCGAAATCA	58645															
QY	657	TTCTGTCTGTGATTTTAAATTTTGAAGGCCGAAAAATAAGTTGTAAAGAGATA	716															
Db	58646	TTCTGTCTGTGATTTTAAATTTTGAAGGCCGAAAAATAAGTTGTAAAGAGATA	58705															
QY	717	AACCCGCCCTATAAATTCATATATTTTCTCCCGCTTTGAAATGTCTCGTTGCTCTCC	776															
Db	58706	AACCCGCCCTATAAATTCATATATTTTCTCTCCCGCTTTGAAATGTCTCGTTGCTCTCC	58765															
QY	777	TCACTTTTCATCAGCCGTTTTTGAATCTCCGCGACTTCGACAGAGAAACAAGGAAGA	836															
Db	58766	TCACTTTTCATCAGCCGTTTTTGAATCTCCGCGACTTCGACAGAGAAACAAGGAAGA	58825															
QY	837	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTGTGAATCTTCTCTCAA	896															
Db	58826	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTGTGAATCTTCTCTCAA	58885															
QY	897	TCTCATCTTCTTCTCCGCTCTTCTTTCCNAGGTAATAGGAACCTTCTGGATCTACTTT	956															
Db	58886	TCTCA---TCTTCTCCGCTCTTCTTTCCNAGGTAATAGGAACCTTCTGGATCTACTTT	58942															
QY	957	ATTGTCTGGATCTCGATCTTGTTTTCTCAATTTCTTTGAGATCTGAAATTCGTTTAAATTT	1016															
Db	58943	ATTGTCTGGATCTCGATCTTGTTTTCTCAATTTCTTTGAGATCTGAAATTCGTTTAAATTT	59002															
QY	1017	GGATCTGTGAACCTCCAAATCTTTTGGTTTTTACTAGAAATCGATCTAAGTTGACCGAT	1076															
Db	59003	GGATCTGTGAACCTCCCAATAATCTTTTGGTTTTTACTAGAAATCGATCTAAGTTGACCGAT	59062															
QY	1077	CAGTTAGCTCGAATATAGCTACAGAAATTTGGCTTTGACCTTGATGGAGAGATCCATGTTC	1136															
Db	59063	CAGTTAGCTCGAATATAGCTACAGAAATTTGGCTTTGACCTTGATGGAGAGATCCATGTTC	59122															
QY	1137	ATGTTACCTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAACTGTTTGAAGTTAGATT	1196															
Db	59123	ATGTTACCTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAACTGTTTGAAGTTAGATT	59182															
QY	1197	GAACTCGAACACTGTCATTTAGATTGAACTGAACTGAACTGTTTAA-GTTAGATGAAGTT	1255															
Db	59183	GAACTCGAACACTGTCATTTAGATTGAACTGAACTGTTTAAAGTTAGATGAAGTT	59242															
QY	1256	TGTTGATAGATTTCTCGAAACCTTAGGATTTCTAGTGTCTGATGTAACAGAAAGCTAT	1315															
Db	59243	TGTTGATAGATTTCTCGAAACCTTTAGGATTTCTAGTGTCTGATGTAACAGAAAGCTAT	59302															
QY	1316	TTCTGATTTCAATCAGGGTTTTTATTTGACTGTATTTGAACTCTTTTTTGTGTTTGCAGTCA	1375															
Db	59303	TTCTGATTTCAATCAGGGTTTTTATTTGACTGTATTTGAACTCTTTTTTGTGTTTGCAGTCA	59362															
QY	1376	TAAAAAATGGCTGAGGCTGACGATTTCAACCAATCGTGTGTGACAACTGGTACTGGAATG	1435															
Db	59363	TAAAAAATGGCTGAGGCTGATGATTTCAACCAATCGTGTGTGACAACTGGTACTGGAATG	59422															
QY	1436	GT 1437																
Db	59423	GT 59424																
RESULT 3																		
AX463618																		
LOCUS	AX463618	6385 bp	DNA	linear	PAT 15-JUL-2002													
DEFINITION	Sequence 15 from Patent W00248335.																	
ACCESSION	AX463618																	
VERSION	AX463618.1	GI:21886378																
KEYWORDS	synthetic construct																	
SOURCE	synthetic construct																	
ORGANISM	other sequences; artificial sequences.																	
REFERENCE	1	Meagher, R.B. and Li, Y.																
AUTHORS	Metal resistant plants and phytoremediation of environmental																	
TITLE	contamination																	
JOURNAL	Patent: WO 0248335-A 15 20-JUN-2002;																	
FEATURES	UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)																	
	Location/Qualifiers																	
	1..6385																	
	/organism="synthetic construct"																	
ORIGIN	/mol_type="unassigned DNA"																	
	/db_xref="taxon:32630"																	
	/note="Plasmid pACT2B-GUS"																	
Query Match 92.9%; Score 1341.8; DB 6; Length 6385;																		
Best Local Similarity 98.8%; Pred. No. 1.8e-279;																		
Matches 1375; Conservative 0; Mismatches 12; Indels 4; Gaps 2;																		
QY	1	ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATATGTAAAG	60															
Db	257	ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATATGTAAAG	316															
QY	61	AAAGTTTTCACGAATATGNNACGACAAAATGGCTACACTCGATGTAAATGGGTATCTCAAC	120															
Db	317	AAAGTTTTCACGAATATGNNACGACAAAATGGCTAGACTCGATGTAAATGGGTATCTCAAC	376															
QY	121	TCACATTTATCTATPACCAACATTTAGTTAGCAAAAATTTAAACAACCTATTTTTATGTAT	180															
Db	377	TCACATTTATCTATPACCAACATTTAGTTAGCAAAAATTTAAACAACCTATTTTTATGTAT	436															
QY	181	GCAAGAGTCAGCATATGTATTAATTTGATTCAGAAATCGTTTTGACGAGTTCCGATGTAGTAG	240															

Db 437 GCAGAGTCAGCATATGTATAATTGATTGAAATCGTTTTGAGGAGTTCGGATGTAGTAG 496  
Qy 241 TAGCCATTATTTAATGTACATACATACTAATCGTGAATAGTATGATGAACACATTGTATCTT 300  
Db 497 TAGCCATTATTTAATGTACATACATACTAATCGTGAATAGTATGATGAACACATTGTATCTT 556  
Qy 301 ATTGTATAAATATCCATAACACATCATGAAGACACITTTCTTTGAGGTCGAAATTAAT 360  
Db 557 ATTGTATAAATATCCATAACACATCATGAAGACACITTTCTTTGAGGTCGAAATTAAT 616  
Qy 361 TATGATACAAATTTCTAATAGAAACGAATTAATTTAGTTGAAATTTGATCAAAATCTAAATG 420  
Db 617 TATGATACAAATTTCTAATAGAAACGAATTAATTTAGTTGAAATTTGATCAAAATCTAAATG 676  
Qy 421 AACAGCCCAACACGACGAGGACTAACGTTGCTGGATTGACTCGGTTTAAAGTTAAACAC 480  
Db 677 AACAGCCCAACACGACGAGGACTAACGTTGCTGGATTGACTCGGTTTAAAGTTAAACAC 736  
Qy 481 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTTCACAGTCATGAAGCCATC 540  
Db 737 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTTCACAGTCATGAAGCCATC 796  
Qy 541 AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAAATTTAGTTTAAAAATTTAGTTAAAC 600  
Db 797 AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAAATTTAGTTTAAAAATTTAGTTAAAC 856  
Qy 601 GAGGGAAGAA - GCTGTCTGACAGCCAGGTTCACGTTATCTTTTACTGTGTGCGAAATGATTC 659  
Db 857 GAGGGAAGAAAGGCTGCTGACAGCCAGGTTCACGTTATCTTTTACTGTGTGCGAAATGATTC 916  
Qy 660 GTGCTGTGCGATTTAAATTTATTTTAAAGGCGGAAATAAGTTGTAAGAGATAAAC 719  
Db 917 GTGCTGTGCGATTTAAATTTATTTTAAAGGCGGAAATAAGTTGTAAGAGATAAAC 976  
Qy 720 CCGCTCTATATAATTCATATATTTTCCCTCCCGCTTTGAAATTTGCTCGTTGCTCTCCCA 779  
Db 977 CCGCTCTATATAATTCATATATTTTCCCTCCCGCTTTGAAATTTGCTCGTTGCTCTCCCA 1036  
Qy 780 CTTTCATCAGCGCTTTTGAATCTCGCGGACTTGACAGAGAAAGAAAGAAAGAACTA 839  
Db 1037 CTTTCATCAGCGCTTTTGAATCTCGCGGACTTGACAGAGAAAGAAAGAAAGAACTA 1096  
Qy 840 AGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGGTTTGAATCTTCTCTCAATCT 899  
Db 1097 AGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGGTTTGAATCTTCTCTCAATCT 1156  
Qy 900 CATCTCTCTTCCGCTCTTCTTCCAGGTAATAGGAACCTTTCTGGATCTACTTTAT 959  
Db 1157 CA---TCITTTCTTCCGCTCTTCTTCCAGGTAATAGGAACCTTTCTGGATCTACTTTAT 1213  
Qy 960 TCCTGGATCTCGATCTTGTCTTCTCAATTTCTTGAGATCTGGAATTCGTTTAAATTTGGA 1019  
Db 1214 TCCTGGATCTCGATCTTGTCTTCTCAATTTCTTGAGATCTGGAATTCGTTTAAATTTGGA 1273  
Qy 1020 TCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGATCAG 1079  
Db 1274 TCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGATCAG 1333  
Qy 1080 TTAGCTCGATATAGCTACAGAAATTTGGCTTGACCTTGATCGAGATCCATGTTGATG 1139  
Db 1334 TTAGCTCGATATAGCTACAGAAATTTGGCTTGACCTTGATCGAGATCCATGTTGATG 1393  
Qy 1140 TTACCTGGGAATGATTTGTATATGTGAATTTGAAATCTGAACTGTTGAAATTTAGATTTGAA 1199  
Db 1394 TTACCTGGGAATGATTTGTATATGTGAATTTGAAATCTGAACTGTTGAAATTTAGATTTGAA 1453  
Qy 1200 TCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTG 1259  
Db 1454 TCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTG 1513  
Qy 1260 TATAGATTTCTCGAAACCTTAGATTTGTAGTGTGTTAGCTTTGAACAGAAAGCTATTCTT 1319  
Db 1514 TATAGATTTCTCGAAACCTTTAGATTTGTAGTGTGTTAGCTTTGAACAGAAAGCTATTCTT 1573

Qy 1320 GATTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTGTGTTTGCAGCTCATAA 1379  
Db 1574 GATTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTGTGTTTGCAGCTCATAA 1633  
Qy 1380 AAATGCGCTGAG 1390  
Db 1634 GGATCCCCGGG 1644  
RESULT 4  
AX463617  
LOCUS AX463617 4526 bp DNA linear PAT 15-JUL-2002  
DEFINITION Sequence 14 from Patent WO0248335.  
ACCESSION AX463617  
VERSION AX463617.1 GI:21886377  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Meagher, R.B. and Li, Y.  
TITLE Metal resistant plants and phytoremediation of environmental  
JOURNAL contamination  
Patent: WO 0248335-A 14 20-JUN-2002;  
UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)  
FEATURES  
source  
1. 4526  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Plasmid pACT2B"  
ORIGIN  
Query Match 92.9%; Score 1341.4; DB 6; Length 4526;  
Best Local Similarity 98.9%; Pred. No. 2.3e-279;  
Matches 1374; Conservative 0; Mismatches 11; Indels 4; Gaps 2;  
Qy 1 ATTATGATCTCAATACATTTGATACATATCTCATCTAGATCTTAGGTATCATATTATGTAAG 60  
Db 257 ATTATGATCTCAATACATTTGATACATATCTCATCTAGATCTTAGGTATCATATTATGTAAG 316  
Qy 61 AAAGTTTGTGCGAATATGNNACGACAAATGGCTACATCTCGATGTAATTTGGTATCTCAAC 120  
Db 317 AAAGTTTGTGCGAATATGNNACGACAAATGGCTACATCTCGATGTAATTTGGTATCTCAAC 376  
Qy 121 TCACATTTATATCTTATACCAACATTTAGTTAGCAAAATTTAAACAACTATTTTATGAT 180  
Db 377 TCACATTTATATCTTATACCAACATTTAGTTAGCAAAATTTAAACAACTATTTTATGAT 436  
Qy 181 GCAAGAGTCAGCATATGTATAATTGATTCAGAAATCGTTTTCAGCAGTTCGGATGTAGTAG 240  
Db 437 GCAAGAGTCAGCATATGTATAATTGATTCAGAAATCGTTTTCAGCAGTTCGGATGTAGTAG 496  
Qy 241 TAGCCATTATTTAATGTACATCTAATCGTGAATAGTATGATGAACAACTTTGATCTTT 300  
Db 497 TAGCCATTATTTAATGTACATCTAATCGTGAATAGTATGATGAACAACTTTGATCTTT 556  
Qy 301 ATTGTATAAATATCCATAACACATCATGAAGACACITTTCTTTGAGGTCGAAATTAAT 360  
Db 557 ATTGTATAAATATCCATAACACATCATGAAGACACITTTCTTTGAGGTCGAAATTAAT 616  
Qy 361 TATGATACAAATTTCTAATAGAAACGAATTTAAATTTAGTTGAAATTTGATCAAAATCTAAATG 420  
Db 617 TATGATACAAATTTCTAATAGAAACGAATTTAAATTTAGTTGAAATTTGATCAAAATCTAAATG 676  
Qy 421 AACAGCCCAACACGACGAGGACTAACCGTTGCTGGATTGACTCGGTTTAAAGTTAAACAC 480  
Db 677 AACAGCCCAACACGACGAGGACTAACCGTTGCTGGATTGACTCGGTTTAAAGTTAAACAC 736  
Qy 481 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTTCACAGTCATGAAGCCATC 540  
Db 737 TAAAAAAGCGAGCTGTCTATGTAAACGCGGATCGAGCAGGTTCACAGTCATGAAGCCATC 796

QY	541	AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAAACAC	600
Db	797	AAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAAACAC	856
QY	601	GAGGGAAAA-GCTGTCTGACAGCCAGGTACAGTTATCTTTTACCTGTGGTCGAAATGATTC	659
Db	857	GAGGGAAAAAGGCTGTCTGACAGCCAGGTACAGTTATCTTTTACCTGTGGTCGAAATGATTC	916
QY	660	GCTGTCTGTCGATTTTAATTAATTTTAAAGGCGGAAAAATAAGTTGTAAAGATAAAC	719
Db	917	GCTGTCTGTCGATTTTAATTAATTTTAAAGGCGGAAAAATAAGTTGTAAAGATAAAC	976
QY	720	CGCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGCTCTCTCA	779
Db	977	CGCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGCTCTCTCA	1036
QY	780	CTTTCATCAGCCGTTTGAATCTCCGGCGACTTGACAGAGAAAGAAACAAGAAAGACTTA	839
Db	1037	CTTTCATCAGCCGTTTGAATCTCCGGCGACTTGACAGAGAAAGAAACAAGAAAGACTTA	1096
QY	840	AGAGAGAAAGTAAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCT	899
Db	1097	AGAGAGAAAGTAAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCT	1156
QY	900	CATCTTCTTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTCTCGGATCTACTTTATT	959
Db	1157	CA---TCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTCTCGGATCTACTTTATT	1213
QY	960	TGCTGGATCTCGATCTTGTGTTTCTCAATTTCTCTGAGATCTGGAATTCGTTTAAATTGGA	1019
Db	1214	TGCTGGATCTCGATCTTGTGTTTCTCAATTTCTCTGAGATCTGGAATTCGTTTAAATTGGA	1273
QY	1020	TCTGTGAACCTCCACTAAATCTTTTGGTTTCTAGAGATCGATCTAAGTTGACCGATCAG	1079
Db	1274	TCTGTGAACCTCCACTAAATCTTTTGGTTTCTAGAGATCGATCTAAGTTGACCGATCAG	1333
QY	1080	TTAGCTCGAATATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCCATGTTTCATG	1139
Db	1334	TTAGCTCGAATATAGCTACAGAAATTTGGCTTGACCTTGATGAGAGATCCCATGTTTCATG	1393
QY	1140	TTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCTGTTTGAAGTTAGATTGAA	1199
Db	1394	TTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCTGTTTGAAGTTAGATTGAA	1453
QY	1200	TCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAAAGTTAGATTGAAGTTTGTG	1259
Db	1454	TCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAAAGTTAGATTGAAGTTTGTG	1513
QY	1260	TATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGATGTTGAACAGAAAGCTATTTCCT	1319
Db	1514	TATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGATGTTGAACAGAAAGCTATTTCCT	1573
QY	1320	GATTCATCAGGGTTTATTTGACTGTATTTGAATCTCTTTTGTGTTGTGTTGAGCTCATAAA	1379
Db	1574	GATTCATCAGGGTTTATTTGACTGTATTTGAATCTCTTTTGTGTTGTGTTGAGCTCATAAA	1633
QY	1380	AAATGGCTG 1388	
Db	1634	GGATCCCGG 1642	
RESULT 5			
AX392814			
LOCUS			
DEFINITION			
Sequence 17 from Patent WO0215701.			
ACCESSION			
AX392814			
VERSION			
AX392814.1 GI:19700916			
KEYWORDS			
synthetic construct			
SOURCE			
synthetic construct			
other sequences; artificial sequences.			
REFERENCE			
1			

AUTHORS	Carozzi,N.B., Rabe,S.M., Miles,P.J., Warren,G.W. and de Haan,P.T.									
TITLE	Novel insecticidal toxins derived from Bacillus thuringiensis insecticidal crystal proteins									
JOURNAL	Patent: WO 0215701-A 17 28-FEB-2002; Syngenta Participations AG (CH)									
FEATURES	Location/Qualifiers									
source	1..11127									
	/organism="synthetic construct"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:32630"									
	/notes="pZU578"									
misc_feature	56..1475									
	/notes="Actin 2 promoter U41998"									
misc_feature	1485..3491									
	/notes="synthetic nucleotide sequence encoding the toxin portion of H04 plus the first 40 amino acids of the CryIAb tail"									
misc_feature	3859..5030									
	/notes="SMAS promoter"									
misc_feature	5052..6271									
	/notes="PMI"									
ORIGIN										
Query Match 90.5%; Score 1306.2; DB 6; Length 11127;										
Best Local Similarity 98.8%; Pred. No. 8.3e-272;										
Matches 1368; Conservative 0; Mismatches 10; Indels 7; Gaps 5;										
QY	1	ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG	60							
Db	93	ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG	152							
QY	61	AAAGTTTTGACGAATATGNNACGACAAATGGCTACACTCGATGTAATTTGGTATCTCAAC	120							
Db	153	AAAGTTTTGACGAATATGGCAGCACAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC	212							
QY	121	TCAACATTATCTTATACCAACATTAGTTAG-CAAAAATTTAAACAACATA-TTTTATATGT	178							
Db	213	TCAACATTATCTTATACCAACATTAGTTAGCAAAAATTTAAACAACATA-TTTTATATGT	272							
QY	179	ATGCAAGAGTCAGCATATGATTAATTTGATTCAGATCGTTTTCACGAGTTTCGGATGTAGT	238							
Db	273	ATGCAAGAGTCAGCATATGATTAATTTGATTCAGATCGTTTTCACGAGTTTCGGATGTAGT	332							
QY	239	AGTAGCCATTATTTAATGTACATACTAATCGTGAATAGTG-ATATGATGAACAACTTGTAT	297							
Db	333	AGTAGCCATTATTTAATGTACATACTAATCGTGAATAGTGAATATGATGAAGCAATTCAT	392							
QY	298	CTTATTGTATAAATATCCATAAACAATCATGATGAAGACACATTTCTTTCAGGGTCTGAATT	357							
Db	393	CTTATTGTATAAATATCCATAAACAATCATGATGAAGACACATTTCTTTCAGGGTCTGAATT	452							
QY	358	AATTATGATACAAATCTTAATAGAAAACGAATTAATTTACGTTGAAATTTGATGAAATCTCAA	417							
Db	453	AATTATGATACAAATCTTAATAGAAAACGAATTAATTTACGTTGAAATTTGATGAAATCTCAA	512							
QY	418	TTGAAACAAGCCCAACACGACGAGGACTAAAGCTTGCCTGGATTGACTCGGTTTAAAGTTAAC	477							
Db	513	TTGAAACAAGCCCAACACGACGAGGACTAAAGCTTGCCTGGATTGACTCGGTTTAAAGTTAAC	572							
QY	478	CACATAAAAAACCGGAGCTGTCTATGTAACAACGCGGATTCAGCAGGTCACAGTCATGAAGCC	537							
Db	573	CACATAAAAAACCGGAGCTGTCTATGTAACAACGCGGATTCAGCAGGTCACAGTCATGAAGCC	632							
QY	538	ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	597							
Db	633	ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTAA	692							
QY	598	CACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACCTGTGGTCGAAATGA	656							
Db	693	CACGAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACCTGTGGTCGAAATGA	752							
QY	657	TTCGTGTCTGTCTGATTTTAAATTTTGTGAAGCGCCGAAAAATAAGTTGTAAGAGATA	716							

```
Db 753 TTCGTGCTGTCGATTTTAATTAATTTTGTAAAGCCGAAATAAAGTTGTAAGATA 812
Qy 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTTGTCGTTGCTCTCC 776
Db 813 AACCCGCTATATAAATTCATATATTTTCTCTCCGCTTTGAATTTGTCGTTGCTCTCC 872
Qy 777 TCACCTTTTCATCAGCGCTTTTGAATCTCCGCGACTTTGACAGAGAAACAAGGAAGA 836
Db 873 TCACCTTTTCATCAGCGCTTTTGAATCTCCGCGACTTTGACAGAGAAACAAGGAAGA 932
Qy 837 CTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCATTTCTCCGTTTGAATCTTCTCTCA 896
Db 933 CTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCATTTCTCCGTTTGAATCTTCTCTCA 992
Qy 897 TCTCATCTTCTTCTCCGCTCTTCTTCTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT 956
Db 993 TCTCA---TCTTCTCCGCTCTTCTTCTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT 1049
Qy 957 ATTTCGTGATCTCGATCTTGTCTTCTCAATTTCTTGGAGATCTGGAATTCGTTTAATTT 1016
Db 1050 ATTTCGTGATCTCGATCTTGTCTTCTCAATTTCTTGGAGATCTGGAATTCGTTTAATTT 1109
Qy 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGATCTAAGTTGACCGAT 1076
Db 1110 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGATCTAAGTTGACCGAT 1169
Qy 1077 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTCACTTGTAGAGAGATCCATGTTTC 1136
Db 1170 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTCACTTGTAGAGAGATCCATGTTTC 1229
Qy 1137 ATGTTACCTCGGAATGATTTGTATATGTAATGGAATCTGAATCTGAATCTGAATCTGAT 1196
Db 1230 ATGTTACCTCGGAATGATTTGTATATGTAATGGAATCTGAATCTGAATCTGAT 1289
Qy 1197 GAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTTGAATTTAGATGAATTT 1256
Db 1290 GAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTTGAATTTAGATGAATTT 1349
Qy 1257 GTGTATAGATTTCTGAAACCTTAGAATTTGATGTTGATGTTGATGTTGATGTTGATGTTG 1316
Db 1350 GTGTATAGATTTCTGAAACCTTTAGGATTTGATGTTGATGTTGATGTTGATGTTGATGTTG 1409
Qy 1317 TCTGATTCATCAGGGTTTATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 1376
Db 1410 TCTGATTCATCAGGGTTTATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 1469
Qy 1377 AAAAA 1381
Db 1470 AAAAA 1474

RESULT 6
LOCUS CQ774496 14184 bp DNA linear PAT 06-MAR-2004
DEFINITION Sequence 58 from Patent WO2004013333.
ACCESSION CQ774496
VERSION CQ774496.1 GI:45237730
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1.
AUTHORS Kock, M., Frank, M. and Badur, R.
TITLE Novel selection method
JOURNAL Patent: WO 2004013333-A 58 12-FEB-2004;
BASF Plant Science GmbH (DE)
FEATURES
source
1..14184
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Beschreibung der kuenstlichen Sequenz: Expression
vector pSUN1-codA-RNA1-Act.-2-Act.A1s-R-ocsf"
```

## ORIGIN

```
Query Match 90.1%; Score 1300.8; DB 6; Length 14184;
Best Local Similarity 98.1%; Pred. No. 1.2e-270;
Matches 1379; Conservative 0; Mismatches 19; Indels 8; Gaps 6;

Qy 1 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG 60
Db 9980 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATTTATGTAAG 10039
Qy 61 AAAGTTTTCACCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
Db 10040 AAAGTTTTCACCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 10099
Qy 121 TCAACATTATATCTTATPACCAAAATTTAGTTAG-CAAAATTTTAAACAACATA-TTTTATGT 178
Db 10100 TCAACATTATATCTTATPACCAAAATTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 10159
Qy 179 ATGCAAGAGTCAAGATATGTAATTAATTTGATTGCAATCGTTTGTGACGAGTTCCGATGTAGT 238
Db 10160 ATGCAAGAGTCAAGATATGTAATTAATTTGATTGCAATCGTTTGTGACGAGTTCCGATGTAGT 10219
Qy 239 AGTAGCCATTATTTAATGTACATATCTAATCGTGAATAGTG-ATATGATGAACATTTGTAT 297
Db 10220 AGTAGCCATTATTTAATGTACATATCTAATCGTGAATAGTG-ATATGATGAACATTTGTAT 10279
Qy 298 CTTATTGTATAAATATCCATAAACAACATCATGAAAGACACATTTTCTTTCAGGGTCTGAAATT 357
Db 10280 CTTATTGTATAAATATCCATAAACAACATCATGAAAGACATTTTCTTTCAGGGTCTGAAATT 10339
Qy 358 AATTATGATACAAATTTCTAATAGAAAACGAATTTAAATTTAGTTGAATTTGATGAATCTTAA 417
Db 10340 AATTATGATACAAATTTCTAATAGAAAACGAATTTAAATTTAGTTGAATTTGATGAATCTTAA 10399
Qy 418 TTGAACAAGCCCAACCAACGAGGAGTAACTGCTGCTGATGACTCGGTTTAAAGTTTAAAC 477
Db 10400 TTGAACAAGCCCAACCAACGAGGAGTAACTGCTGCTGATGACTCGGTTTAAAGTTTAAAC 10459
Qy 478 CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTCACAGTCAATGAAGCC 537
Db 10460 CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTCACAGTCAATGAAGCC 10519
Qy 538 ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTTAA 597
Db 10520 ATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTTAGTTTAA 10579
Qy 598 CACGAGGGGAAA-GCTGTCTGACACGCGGATCACTGTTTCTGCTGCTGCTGCTGCTGCTGCT 656
Db 10580 CACGAGGGGAAAAGGCTGTCTGACAGCCAGGTCACGTTTATCTTCTGCTGCTGCTGCTGCT 10639
Qy 657 TTCGTGCTGTGCGATTTTAAATTTTAAAGGCGCGAAATAAAGTTGTAAGAGATA 716
Db 10640 TTCGTGCTGTGCGATTTTAAATTTTAAAGGCGCGAAATAAAGTTGTAAGAGATA 10699
Qy 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTTGTCGTTGCTCTCC 776
Db 10700 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTTGTCGTTGCTCTCC 10759
Qy 777 TCACCTTTTCATCAGCGCTTTTGAATCTCCGCGACTTTGACAGAGAAACAAGGAAGA 836
Db 10760 TCACCTTTTCATCAGCGCTTTTGAATCTCCGCGACTTTGACAGAGAAACAAGGAAGA 10819
Qy 837 CTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCATTTCTCCGTTTGAATCTTCTCTCA 896
Db 10820 CTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCATTTCTCCGTTTGAATCTTCTCTCA 10879
Qy 897 TCTCATCTTCTTCTCCGCTCTTCTTCTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT 956
Db 10880 TCTCA---TCTTCTTCTCCGCTCTTCTTCTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT 10936
Qy 957 ATTTCGTGATCTCGATCTTGTCTTCTCAATTTCTTGGAGATCTGGAATTCGTTTAATTT 1016
Db 10937 ATTTCGTGATCTCGATCTTGTCTTCTCAATTTCTTGGAGATCTGGAATTCGTTTAATTT 10996
```



Qy	1017	GGATCTGTGAACTCCACCTAAATCTTTTGGTTTACTAGAACTGATCTAAAGTTGACCGAT	1076
Db	10997	GGATCTGTGAACTCCACCTAAATCTTTTGGTTTACTAGAACTGATCTAAAGTTGACCGAT	11056
Qy	1077	CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGTGAGAGATCCATGTTTC	1136
Db	11057	CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGTGAGAGATCCATGTTTC	11116
Qy	1137	ATGTTACCTGGGAAATGATTTGTATATGTAATTTGAAATCTGAACTGTTGAAAGTTAGATT	1196
Db	11117	ATGTTACCTGGGAAATGATTTGTATATGTAATTTGAAATCTGAACTGTTGAAAGTTAGATT	11176
Qy	1197	GAACTCGAACACTGTCTCAATGTTAGATTGAATCTGAACTGTTTAA-GTTAGATGAAAGTT	1255
Db	11177	GAACTCGAACACTGTCTCAATGTTAGATTGAATCTGAACTGTTTAAAGGTTAGATGAAAGTT	11236
Qy	1256	TGTGTATAGATCTTTCGAAACCTTTAGGATTTGTAGTGTGCTGATGTAACAGAAAGCTAT	1315
Db	11237	TGTGTATAGATCTTTCGAAACCTTTAGGATTTGTAGTGTGCTGATGTAACAGAAAGCTAT	11296
Qy	1316	TTCTGATTCGAATCAGGGTTTATTTGACTGTATTTGAACTCTTTTGTGTGTTTGCAGCTCA	1375
Db	11297	TTCTGATTCGAATCAGGGTTTATTTGACTGTATTTGAACTCTTTTGTGTGTTTGCAGCTCA	11356
Qy	1376	TAAAAAATGGCTGAGGCTGACGATAT 1401	
Db	11357	TAAAAAATGGCTGAGGCTGACGATAT 11382	
RESULT 7			
AX463620	AX463620	3450 bp	DNA linear PAT 15-JUL-2002
LOCUS			
DEFINITION	Sequence 17 from Patent WO0248335.		
ACCESSION	AX463620		
VERSION	AX463620.1 GI:21886380		
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Meagher,R.B. and Li,Y.		
TITLE	Metal resistant plants and phytoremediation of environmental contamination		
JOURNAL	Patent: WO 0248335-A 17 20-JUN-2002;		
FEATURES	UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)		
source	Location/Qualifiers		
	1..3450		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Plasmid pECSACT2"		
ORIGIN			
Query Match	89.9%;	Score 1298.4;	DB 6; Length 3450;
Best Local Similarity	98.1%;	Pred.No. 4.6e-270;	
Matches 1376;	Conservative 0;	Mismatches 18;	Indels 8; Gaps 6;
Qy	1	ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATTTGTAAG	60
Db	9	ATGCTGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTTATCATTTGTAAG	68
Qy	61	AAAGTTTTCAGCAATATGNNACGACAAATGGCTACACTCGATGTAATTTGGTATCTCAAC	120
Db	69	AAAGTTTTCAGCAATATGGCAGACAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC	128
Qy	121	TCAACATTATCTTATACCAACATTAGTTAG-CAAAATTTTAAACAACCTA-TTTTATGCT	178
Db	129	TCAACATTATCTTATACCAACATTAGTTAGCAGAAATTTTAAACAACCTATTTTATGCT	188
Qy	179	ATGCAAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTTCAGAGTTCCGATGTAGT	238
Db	189	ATGCAAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTTCAGAGTTCCGATGTAGT	248

QY	239	AGTAGCCATTATTTAAATGTACATACTAATCGTGAATAGTG-ATATGATGAAACATTGTAT	297
Db	249	AGTAGCCATTATTTAAATGTACATACTAATCGTGAATAGTGATATGATGAACATTGTAT	308
QY	298	CTTATTGTATAAATATCCATAAACAATCATGAAAGACACATTTCTTTTCAGGGTCTGAATT	357
Db	309	CTTATTGTATAAATATCCATAAACAATCATGAAAGACACATTTCTTTTCAGGGTCTGAATT	368
QY	358	AAATTATGATACAAATTTCTTAATAGAAAACGAAATTAATTTAGTTGAATTTGATGAAATCTAA	417
Db	369	AAATTATGATACAAATTTCTTAATAGAAAACGAAATTAATTTAGTTGAATTTGATGAAATCTAA	428
QY	418	TTGAACAAGCCCAACACGACGAGGACTAAGTTGCTCGATTGACTTCGGTTTAAAGTTAAC	477
Db	429	TTGAACAAGCCCAACACGACGAGGACTAAGTTGCTCGATTGACTTCGGTTTAAAGTTAAC	488
QY	478	CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTACAGTCAATGAAGCC	537
Db	489	CACTAAAAAAGCGAGCTGTCTATGTAAACAACGCGGATCGAGCAGGTACAGTCAATGAAGCC	548
QY	538	ATCAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTAGTTAA	597
Db	549	ATCAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTAGTTAA	608
QY	598	CACGAGGGGAAAA-GCTGTCTGACAGCCAGGTCACTGTTATCTTTTACCTGTGGTCGAAATGA	656
Db	609	CACGAGGGGAAAAAGGCTGTCTGACAGCCAGGTCACTGTTATCTTTTACCTGTGGTCGAAATGA	668
QY	657	TTCGTGTCTGTCGATTTTAAATTAATTTTTTGAAGGCCGAAAAATAAAGTTGTAAGAGATA	716
Db	669	TTCGTGTCTGTCGATTTTAAATTAATTTTTTGAAGGCCGAAAAATAAAGTTGTAAGAGATA	728
QY	717	AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTTGTCTCTCC	776
Db	729	AACCCGCTATATAAATTTTCTCTCCCGCTTTTGAATTTGTCTCGTTTGTCTCTCTCC	788
QY	777	TCACTTTTCATCAGCGTTTGTGAATCTCCGGGACCTTGACAGAGAAAGAACAGAAAGA	836
Db	789	TCACTTTTCATCAGCGCTTTTGAATCTCCGGGACCTTGACAGAGAAAGAACAGAAAGA	848
QY	837	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCTCAA	896
Db	849	CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCTCAA	908
QY	897	TCATCTTCTTCTTCCGCTCTTTCTTTCCAAAGGTAATAGGAACTTTCTCGAATCTACTTT	956
Db	909	TCCTCA---TCTTCTTCCGCTCTTTCTTTCCAAAGGTAATAGGAACTTTCTCGAATCTACTTT	965
QY	957	ATTTCTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTGGAGATCTGAAATTCGTTTAAATTT	1016
Db	966	ATTTCTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTGGAGATCTGAAATTCGTTTAAATTT	1025
QY	1017	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAACTGATCTAAGTTGACCGAT	1076
Db	1026	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAACTGATCTAAGTTGACCGAT	1085
QY	1077	CAGTTAGCTCGATTATAGCTACAGAAATTTGCGTTTGAACCTTGCATGAGAGATCCATGTTTC	1136
Db	1086	CAGTTAGCTCGATTATAGCTACAGAAATTTGCGTTTGAACCTTGCATGAGAGATCCATGTTTC	1145
QY	1137	ATGTTTACCTGGGAAATGATTTGTATATGTAATTCGAAATCTGAACTGTTTGAAGTTAGATT	1196
Db	1146	ATGTTTACCTGGGAAATGATTTGTATATGTAATTCGAAATCTGAACTGTTTGAAGTTAGATT	1205
QY	1197	GAACTCGAACCTGTCAATGTTTAGATTGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT	1255
Db	1206	GAACTCGAACCTGTCAATGTTTAGATTGAATCTGAACACTGTTTAAAGGTTAGATGAAGTT	1265
QY	1256	TGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGATGTAACAGAAAGCTAT	1315
Db	1266	TGTGTATAGATTTCTCGAAACCTTTAGGATTTGTAGTGTGCTGATGTAACAGAAAGCTAT	1325



QY 1316 TTCTGATCAATCAGGGTTTATTTGACTGATGAACTCTTTTGTGTGTTTGCAGCTCA 1375  
Db 1326 TTCTGATCAATCAGGGTTTATTTGACTGATGAACTCTTTTGTGTGTTTGCAGCTCA 1385  
QY 1376 TAAATAATGGCTGAGGCTGACG 1397  
Db 1386 TAAACNCGGAATCCCGAGC 1407

RESULT 8  
LOCUS CQ814616 7474 bp DNA linear PAT 24-MAY-2004  
DEFINITION Sequence 7 from Patent WO2004039986.  
ACCESSION CQ814616  
VERSION CQ814616.1 GI:47603799  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Ellis,D.M., Negrotto,D.V., Shi,L., Shotkoski,F.A. and Thomas,C.R.  
TITLE Cot102 insecticidal cotton  
JOURNAL Patent: WO 2004039986-A 7 13-MAY-2004;  
Syngenta Participations AG (CH)  
FEATURES  
source location/Qualifiers  
1.7474  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Cot102 nucleotide motif"

ORIGIN  
Query Match 89.9%; Score 1298.4; DB 6; Length 7474;  
Best Local Similarity 98.8%; Pred. No. 4.2e-270;  
Matches 1370; Conservative 0; Mismatches 8; Indels 8; Gaps 6;  
QY 1 ATTATGATCTCAATATACATATGATATCATCTCATCTAGATCTAGGTATCATATGTAAG 60  
Db 3228 ATTATGATCTCAATATACATATGATATCATCTCATCTAGATCTAGGTATCATATGTAAG 3287  
QY 61 AAAGTTTGGACCAATATGNNACGACAAATGCTACACTCGATGTAATGGTATCTCAAC 120  
Db 3288 AAAGTTTGGACCAATATGNNACGACAAATGCTACACTCGATGTAATGGTATCTCAAC 3347  
QY 121 TCAACATTATATCTTATACCAACATTAGTTAG - CAAAATTTAAACAACTA - TTTTATATGT 178  
Db 3348 TCAACATTATATCTTATACCAACATTAGTTAGCAAAATTTAAACAACTATTTTATGT 3407  
QY 179 ATGCAAGAGTCAGCATATGATATTAATGATTCAGATTCGGTTTGGACGAGTTCGGATGTAGT 238  
Db 3408 ATGCAAGAGTCAGCATATGATATTAATGATTCAGATTCGGTTTGGACGAGTTCGGATGTAGT 3467  
QY 239 AGTAGCCATTATTTAATGTACATCTAATTCGGTGAATAGTG - ATATGATGAAACATGTAT 297  
Db 3468 AGTAGCCATTATTTAATGTACATCTAATTCGGTGAATAGTGAAATATGATGAAACATGTAT 3527  
QY 298 CTATTGTATATAATATCCATAACACATCATCAAGACACTTTCTTTCAGGGTCTGAAAT 357  
Db 3528 CTATTGTATATAATATCCATAACACATCATCAAGACACTTTCTTTCAGGGTCTGAAAT 3587  
QY 358 AATTATGATACAAATTTCTAATAGAAAACGAATTAATAGTTGAAATGATGAAATCTAA 417  
Db 3588 AATTATGATACAAATTTCTAATAGAAAACGAATTAATAGTTGAAATGATGAAATCTAA 3647  
QY 418 TTGAACAAGCCAAACACGACGAGGACTAACGTTGCTGATGATGATTCGGTTTAAAGTTAAC 477  
Db 3648 TTGAACAAGCCAAACACGACGAGGACTAACGTTGCTGATGATGATTCGGTTTAAAGTTAAC 3707  
QY 478 CACTAAAAAAGCGAGCTGTCTATGTAACACGCGGATTCGAGCAGGTACAGTTCATGAAGCC 537  
Db 3708 CACTAAAAAAGCGAGCTGTCTATGTAACACGCGGATTCGAGCAGGTACAGTTCATGAAGCC 3767  
QY 538 ATCAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTAGTTAA 597

Db 3768 ATCAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTAGTTAA 3827  
QY 598 CACGAGGAGAAA - GCTGTCTGACACGACAGGTACAGTTATCTTTTACCTGTGGTCCGAATCA 656  
Db 3828 CACGAGGAGAAAAGGCTGTGACACGACAGGTACAGTTATCTTTTACCTGTGGTCCGAATCA 3887  
QY 657 TTCGTGTCTGTCGATTTTAAATTTATTTTGAAGGCCGAAATAAAGTTGTAAGAGATA 716  
Db 3888 TTCGTGTCTGTCGATTTTAAATTTATTTTGAAGGCCGAAATAAAGTTGTAAGAGATA 3947  
QY 717 AACCCGCTATATAAATTTATATATTTTCTCCCGCTTTGAAATTTGTCGTTGTCCTCC 776  
Db 3948 AACCCGCTATATAAATTTATATATTTTCTCCCGCTTTGAAATTTGTCGTTGTCCTCC 4007  
QY 777 TCACCTTTCATCAGCGCTTTTGAATCTCCGGGACCTTTGACAGAGAAACAAAGGAAGA 836  
Db 4008 TCACCTTTCATCAGCGCTTTTGAATCTCCGGGACCTTTGACAGAGAAACAAAGGAAGA 4067  
QY 837 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTTGAATCTTCTCTCAA 896  
Db 4068 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTTGAATCTTCTCTCAA 4127  
QY 897 TCTCATCTTCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAACTTTCTGGAATCTACTTT 956  
Db 4128 TCTCA - - - TCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAACTTTCTGGAATCTACTTT 4184  
QY 957 ATTTGCTGATCTCGATCTGTTTCTCAATTTCTGAGATCTGGAATTCGTTTAAATTT 1016  
Db 4185 ATTTGCTGATCTCGATCTGTTTCTCAATTTCTGAGATCTGGAATTCGTTTAAATTT 4244  
QY 1017 GGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAACTGATCTAAGTTGACCGAT 1076  
Db 4245 GGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAACTGATCTAAGTTGACCGAT 4304  
QY 1077 CAGTTAGTCTGATATAGCTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTC 1136  
Db 4305 CAGTTAGTCTGATATAGCTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTC 4364  
QY 1137 ATGTTACCTGGGAATGATTTGTATATGTAATCTGAAATCTGAACTGTTGAAGTTAGATT 1196  
Db 4365 ATGTTACCTGGGAATGATTTGTATATGTAATCTGAAATCTGAACTGTTGAAGTTAGATT 4424  
QY 1197 GAATCTGAACACTGTCAATGTTAGATGTAATCTGAACACTGTTTAA - GTTAGATGAAGTT 1255  
Db 4425 GAATCTGAACACTGTCAATGTTAGATGTAATCTGAACACTGTTTAAAGTTAGATGAAGTT 4484  
QY 1256 TGTGTATAGATCTTCGAAACCTTAGGATTTGTAGTGTGCTACGTTGAACAGAAAGCTAT 1315  
Db 4485 TGTGTATAGATCTTCGAAACCTTAGGATTTGTAGTGTGCTACGTTGAACAGAAAGCTAT 4544  
QY 1316 TTCTGATTCATTCAGGGTTTATTTGACCTGATTTGAATCTTTTGTGTGTTTGCAGCTCA 1375  
Db 4545 TTCTGATTCATTCAGGGTTTATTTGACCTGATTTGAATCTTTTGTGTGTTTGCAGCTCA 4604  
QY 1376 TAAAAA 1381  
Db 4605 TAAAAA 4610

RESULT 9  
LOCUS CQ814630 9356 bp DNA linear PAT 24-MAY-2004  
DEFINITION Sequence 21 from Patent WO2004039986.  
ACCESSION CQ814630  
VERSION CQ814630.1 GI:47603812  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Ellis,D.M., Negrotto,D.V., Shi,L., Shotkoski,F.A. and Thomas,C.R.  
TITLE Cot102 insecticidal cotton

JOURNAL Patent: WO 2004039986-A 21 13-MAY-2004;  
Syngenta Participations AG (CH)  
FEATURES Location/Qualifiers  
source 1..9356  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="COT102 nucleotide motif"

ORIGIN

Query Match 89.9%; Score 1298.4; DB 6; Length 9356;  
Best Local Similarity 98.8%; Pred. No. 4.1e-270;  
Matches 1370; Conservative 0; Mismatches 8; Indels 8; Gaps 6;  
QY 1 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 60  
DB 4614 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 4673  
QY 61 AAAGTTTGGACGAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120  
DB 4674 AAAGTTTGGACGAATATGCGACGACAAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC 4733  
QY 121 TCAACATATATCTTATACAAACATTTAGTTAG-CAAAATTTAAACAACTA-TTTTATGT 178  
DB 4734 TCAACATATATCTTATACAAACATTTAGTTAGCAAAAATTTAAACAACTATTTTATGT 4793  
QY 179 ATGCAAGAGTCAGCATATGTAATTTGATTCAGAACTGTTTGGAGGTTCCGATGTAGT 238  
DB 4794 ATGCAAGAGTCAGCATATGTAATTTGATTCAGAACTGTTTGGAGGTTCCGATGTAGT 4853  
QY 239 AGTAGGCATTTTAAATGTACATCTAATTCGTGAATAGTG-ATATGATGAAACATTTGTAT 297  
DB 4854 AGTAGGCATTTTAAATGTACATCTAATTCGTGAATAGTGATGATGATGAAACATTTGTAT 4913  
QY 298 CTATTGTATAAATATCCATAAACACATCATGAAAGACACTTTCTTTCAGGTCGTGAAT 357  
DB 4914 CTATTGTATAAATATCCATAAACACATCATGAAAGACACTTTCTTTCAGGTCGTGAAT 4973  
QY 358 AATTATGATACAAATCTATAGAAACGAATTAATTTAGTTGAAATGTATGAAATCTAA 417  
DB 4974 AATTATGATACAAATCTATAGAAACGAATTAATTTAGTTGAAATGTATGAAATCTAA 5033  
QY 418 TTGAACAACGCCAACACGACGAGGACTAACGTTGCTCGATTTGACTCGGTTTAAAGTTAAC 477  
DB 5034 TTGAACAACGCCAACACGACGAGGACTAACGTTGCTCGATTTGACTCGGTTTAAAGTTAAC 5093  
QY 478 CACTAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCGAGCAGGTCAAGTATGAAGCC 537  
DB 5094 CACTAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCGAGCAGGTCAAGTATGAAGCC 5153  
QY 538 ATCAAGCAAAAGAACTAATCCAAGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 597  
DB 5154 ATCAAGCAAAAGAACTAATCCAAGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 5213  
QY 598 CACGAGGGAATAA-GCTGTCTGACAGCCAGGTCACTTATCTTTTACCTGTGGTCGAAATCA 656  
DB 5214 CACGAGGGAATAAAGGCTGTCTGACAGCCAGGTCACTTATCTTTTACCTGTGGTCGAAATCA 5273  
QY 657 TTCTGTCTGTCTGATTTTAAATTTATTTTAAAGGCCGAAATAAAGTTGTAAGAGATA 716  
DB 5274 TTCTGTCTGTCTGATTTTAAATTTATTTTAAAGGCCGAAATAAAGTTGTAAGAGATA 5333  
QY 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATTTCTCGTTGCTCTCC 776  
DB 5334 AACCCGCTATATAAATTCATATATTTTCTCTCCCGCTTTGAAATTTCTCGTTGCTCTCC 5393  
QY 777 TCACCTTTTCATCAGCCGTTTGAATCTCCGCGACTTTCAGAGAGAAACAAGGAAGAAGA 836  
DB 5394 TCACCTTTTCATCAGCCGTTTGAATCTCCGCGACTTTCAGAGAGAAACAAGGAAGAAGA 5453  
QY 837 CTAAGAGAGAAAGTAAGAGATAATCAGAGATTCATTTCTCGTTTGAATTTCTTCTCAA 896  
DB 5454 CTAAGAGAGAAAGTAAGAGATAATCAGAGATTCATTTCTCGTTTGAATTTCTTCTCAA 5513

QY 897 TCTCATCTTCTTCTCCGCTCTTTCTTTCCAAGGTAATAGGAACCTTTCTGATCTACTTT 956  
DB 5514 TCTCA--TCTTCTTCCGCTCTTTCTTTCCAAGGTAATAGGAACCTTTCTGATCTACTTT 5570  
QY 957 ATTTCCTGGAATCTCGATCTTGTGTTTCTCAATTTCTTGGAGATCTGGAATCTGTTAAATTT 1016  
DB 5571 ATTTCCTGGAATCTCGATCTTGTGTTTCTCAATTTCTTGGAGATCTGGAATCTGTTAAATTT 5630  
QY 1017 GGATCTGTGAACCTCCACTTAATCTTTTGGTTTACTAGAACGATCTAAGTTGACCGAT 1076  
DB 5631 GGATCTGTGAACCTCCACTTAATCTTTTGGTTTACTAGAACGATCTAAGTTGACCGAT 5690  
QY 1077 CAGTTAGTCTCGAATATAGCTACCAAAATTTGGCTTGAACCTTATGATGGAGAGATCCCATGTTTC 1136  
DB 5691 CAGTTAGTCTCGAATATAGCTACCAAAATTTGGCTTGAACCTTATGATGGAGAGATCCCATGTTTC 5750  
QY 1137 ATGTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACTGTTGAAAGTTAGATT 1196  
DB 5751 ATGTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACTGTTGAAAGTTAGATT 5810  
QY 1197 GAATCTGAACACTGTCAATGTAGATTGAATCTGAACACTGTTTAAAGTTAGTAGATT 1255  
DB 5811 GAATCTGAACACTGTCAATGTAGATTGAATCTGAACACTGTTTAAAGTTAGTAGATT 5870  
QY 1256 TGTGTATAGATTCTTTCGAAACCTTAGGATTTGTAGTTCGTACGTTGAACAGAAAGCTAT 1315  
DB 5871 TGTGTATAGATTCTTTCGAAACCTTAGGATTTGTAGTTCGTACGTTGAACAGAAAGCTAT 5930  
QY 1316 TTCTGATTAATTCAGGTTTATTTGACTGTATGTAATCTTTTGTGTTGTTGAGCTCA 1375  
DB 5931 TTCTGATTAATTCAGGTTTATTTGACTGTATGTAATCTTTTGTGTTGTTGAGCTCA 5990  
QY 1376 TAAAAA 1381  
DB 5991 TAAAAA 5996  
RESULT 10  
AX463619  
LOCUS AX463619  
DEFINITION Sequence 16 from Patent WO0248335.  
ACCESSION AX463619  
VERSION AX463619.1 GI:21886379  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Meagher,R.B. and Li,Y.  
TITLE Metal resistant plants and phytoremediation of environmental  
JOURNAL contamination  
Patent: WO 0248335-A 16 20-JUN-2002;  
UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)  
FEATURES Location/Qualifiers  
source 1..3408  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Plasmid pAtPCSACT2"  
ORIGIN  
Query Match 89.8%; Score 1297.4; DB 6; Length 3408;  
Best Local Similarity 98.5%; Pred. No. 7.6e-270;  
Matches 1372; Conservative 0; Mismatches 13; Indels 8; Gaps 6;  
QY 1 ATTATGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 60  
DB 9 ATGCTGATCTCAAAATACATTGATACATATCTCATCTAGATCTAGGTATCATATTGTAAAG 68  
QY 61 AAAGTTTGGACGAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120  
DB 69 AAAGTTTGGACGAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 128

```
QY 121 TCAACATTATATCTTATACCAACATTAGTTAG-CAAAAATTTAAACAACCTA-TTTTATGT 178
Db 129 TCAACATTATATCTTATACCAACATTAGTTAGCAAAAATTTAAACAACCTATTTTATGT 188
QY 179 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAACTCGTTCGAGTTCGGATGTAGT 238
Db 189 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAACTCGTTCGAGTTCGGATGTAGT 248
QY 239 AGTAGCCATTATTTAATGTACATACTAATCGTGAATAGTG-ATATGATGAAACATTGTAT 297
Db 249 AGTAGCCATTATTTAATGTACATACTAATCGTGAATAGTGATGATGATGATGATGAT 308
QY 298 CTTATTTGTATTAATATCCATAACACATCATGAAAGACACTTTCTTTTCAGGTCCTGAATT 357
Db 309 CTTATTTGTATTAATATCCATAACACATCATGAAAGACACTTTCTTTTCAGGTCCTGAATT 368
QY 358 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAAATCTAA 417
Db 369 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAAATCTAA 428
QY 418 TTGAACAAGCCCAACACGACGAGGACTAAGCTTTCCTGGAATGACTTCGGTTTAAAGTTAAC 477
Db 429 TTGAACAAGCCCAACACGACGAGGACTAAGCTTTCCTGGAATGACTTCGGTTTAAAGTTAAC 488
QY 478 CACTAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCGAGCGAGTCAAGTCAATGAAGCC 537
Db 489 CACTAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCGAGCGAGTCAAGTCAATGAAGCC 548
QY 538 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAA 597
Db 549 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAA 608
QY 598 CACGAGGGAAAA-GCTGTCTGACAGCGAGTCAAGTTCCTTACCTGTGGTCGAAATGA 656
Db 609 CACGAGGGAAAAAGGCTGTCTGACAGCGAGTCAAGTTCCTTACCTGTGGTCGAAATGA 668
QY 657 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGGCGGAAAAATAAAGTTGTAAGAGATA 716
Db 669 TTCGTGTCTGTGATTTTAAATTTATTTTAAAGGCGGAAAAATAAAGTTGTAAGAGATA 728
QY 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGATTTGCTCGTTGCTCC 776
Db 729 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGATTTGCTCGTTGCTCC 788
QY 777 TCACCTTTTCATCAGCCGTTTTGAACTCTCCGCGACTTGACAGAGAAGAAACAAGGAAGA 836
Db 789 TCACCTTTTCATCAGCCGTTTTGAACTCTCCGCGACTTGACAGAGAAGAAACAAGGAAGA 848
QY 837 CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTTGAATCTTCTCAA 896
Db 849 CTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTTGAATCTTCTCAA 908
QY 897 TCCTCATCTTCTTCTCCGCTCTTTCTTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT 956
Db 909 TCTCA--TCTTCTTCTCCGCTCTTTCTTTCCAGGTAATAGGAACCTTTCTGGATCTACTTT 965
QY 957 ATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTTGAGATCTCGAATTCGTTTAAATTT 1016
Db 966 ATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTTGAGATCTCGAATTCGTTTAAATTT 1025
QY 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAAATCTAAGTTGACCGAT 1076
Db 1026 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAAATCTAAGTTGACCGAT 1085
QY 1077 CAGTTAGCTCGATATAGTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTC 1136
Db 1086 CAGTTAGCTCGATATAGTACAGAAATTTGGCTTGACCTTGATCGAGAGATCCATGTC 1145
QY 1137 ATGTTACCTGGAAATGATTTGTATGTGAATTCGAAATCTGAACTGTTGAGTTAGATT 1196
Db 1146 ATGTTACCTGGAAATGATTTGTATGTGAATTCGAAATCTGAACTGTTGAGTTAGATT 1205
```

```
QY 1197 GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTGTTAA-GTTAGATGAAGTT 1255
Db 1206 GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTGTTAAAGGTTAGATGAAGTT 1265
QY 1256 TGTGTATAGATTCTTCGAAACCTTACGATTTCTAGTGTCTGAGTTCGAGTTCGAGAAAGCTAT 1315
Db 1266 TGTGTATAGATTCTTCGAAACCTTACGATTTCTAGTGTCTGAGTTCGAGTTCGAGAAAGCTAT 1325
QY 1316 TTCTGATTCAAATCAGGCTTTTATTTGACTGTATTTGAACCTCTTTTGTGTGTTTGCAGCTCA 1375
Db 1326 TTCTGATTCAAATCAGGCTTTTATTTGACTGTATTTGAACCTCTTTTGTGTGTTTGCAGCTCA 1385
QY 1376 TAAAAAATGGCTG 1388
Db 1386 TAAACCATGGCAG 1398

RESULT 11
AX463621 LOCUS 2857 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 18 from Patent WO0248335.
ACCESSION AX463621
VERSION AX463621.1 GI:21886381
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Meagher, R.B. and Li, Y.
TITLE Metal resistant plants and phytoremediation of environmental
JOURNAL contamination
PATENT: WO 0248335-A 18 20-JUN-2002;
UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)
FEATURES
LOCATION/Qualifiers
1..2857
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid pGSACT2"

ORIGIN

Query Match 89.8%; Score 1297; DB 6; Length 2857;
Best Local Similarity 98.6%; Pred. No. 9.5e-270;
Matches 1371; Conservative 0; Mismatches 12; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGCTTATCATTTATGTAG 60
Db 9 ATGCTGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGCTTATCATTTATGTAG 68
QY 61 AAAAGTTTTCGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
Db 69 AAAAGTTTTCGCAATATGNNACGACAAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC 128
QY 121 TCAACATTATATCTTATACCAACATTAGTTAG-CAAAAATTTAAACAACCTA-TTTTATGT 178
Db 129 TCAACATTATATCTTATACCAACATTAGTTAGCAAAAATTTAAACAACCTATTTTATGT 188
QY 179 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAACTCGTTCGAGTTCGGATGTAGT 238
Db 189 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAACTCGTTCGAGTTCGGATGTAGT 248
QY 239 AGTAGCCATTATTTAATGTACATACTAATCGTGAATAGTG-ATATGATGAAACATTGTAT 297
Db 249 AGTAGCCATTATTTAATGTACATACTAATCGTGAATAGTGATGATGATGATGATGAT 308
QY 298 CTTATTTGTATTAATATCCATAACACATCATGAAAGACACTTTCTTTTCAGGTCCTGAATT 357
Db 309 CTTATTTGTATTAATATCCATAACACATCATGAAAGACACTTTCTTTTCAGGTCCTGAATT 368
QY 358 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAAATCTAA 417
Db 369 AATTATGATACAAATCTTAATAGAAAACGAATTAATTTAGCTTGAATTTGATGAAATCTAA 428
```



Db 301 CGGTTTAAGTTAACCACTAAAAACCGAGCTGTCATGTAAACACCGCGATCGAGCAGGTC 360  
Qy 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTT 583  
Db 361 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTT 420  
Qy 584 TAAAAATTAGTTAAACAGAGGGAAGAA- GCTGCTGACAGCGAGGTACAGTTATCTTTACC 642  
Db 421 TAAAAATTAGTTAAACAGAGGGAAGAAAGGCTGCTGACAGCGAGGTACAGTTATCTTTACC 480  
Qy 643 TGTGGTCGAAATGATTCGTGCTGCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAA 702  
Db 481 TGTGGTCGAAATGATTCGTGCTGCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAA 540  
Qy 703 AGTTGTAAGAGATAAACCCGCCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG 762  
Db 541 AGTTGTAAGAGATAAACCCGCCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG 600  
Qy 763 TCTCGTTGTCCTCCTCCTCCTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAG 822  
Db 601 TCTCGTTGTCCTCCTCCTCCTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAG 660  
Qy 823 AACAGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTT 882  
Db 661 AACAGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTT 720  
Qy 883 TGAATCTTCTCCTCAATCTCATCTTCTTCTCCGCTCTTCTTCCAGGTAATAGGAACTT 942  
Db 721 TGAATCTTCTCCTCAATCTCA- --TCTTCTCCGCTCTTCTTCCAGGTAATAGGAACTT 777  
Qy 943 TCTGGATCTACTTTATTTGCTGGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTGG 1002  
Db 778 TCTGGATCTACTTTATTTGCTGGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTGG 837  
Qy 1003 AATTCTGTTTAAATTTGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGAT 1062  
Db 838 AATTCTGTTTAAATTTGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGAT 897  
Qy 1063 CTAAGTTGACCGATCAGTTAGCTCGATATAGCTACCCAGAAATTTGGCTTGGATCTGG 1122  
Db 898 CTAAGTTGACCGATCAGTTAGCTCGATATAGCTACCCAGAAATTTGGCTTGGATCTGG 957  
Qy 1123 AGAGATCCATGTTTATGTTTACCTGGGAATGATTTGTATATGTGAATGAAATCTGAAT 1182  
Db 958 AGAGATCCATGTTTATGTTTACCTGGGAATGATTTGTATATGTGAATGAAATCTGAAT 1017  
Qy 1183 GTTGAAGTTAGATTTGAATCTGAACCTGCAATGTTAGATTTGAATCTGAACCTGTTAA 1242  
Db 1018 GTTGAAGTTAGATTTGAATCTGAACCTGCAATGTTAGATTTGAATCTGAACCTGTTAA 1077  
Qy 1243 GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGTTACGTTG 1302  
Db 1078 GTTAGATGAAGTTTGTGTATAGATTTCTCGAACTTTAGGATTTGTAGTGTGTTACGTTG 1137  
Qy 1303 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAATCTTTTGTG 1362  
Db 1138 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAATCTTTTGTG 1197  
Qy 1363 TGTGTTGAGCTCATAAAGAAAGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1422  
Db 1198 TGTGTTGAGCTCATAAAGAAAGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1257  
Qy 1423 TGGTACTGGAATGGT 1437  
Db 1258 TGGTACCGGTATGGT 1272

RESULT 13  
AX461227  
LOCUS AX461227 1217 bp DNA linear PAT 08-JUL-2002  
DEFINITION Sequence 156 from Patent W00198480.  
ACCESSION AX461227  
VERSION AX461227.1 GI:21726435

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

SpERMATOPHYTES; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; ArabidopsiS.

1

REFERENCE

AUTHORS

Budworth, P., Brown, D., Chang, H.S., Zhu, T., Han, B., Wang, X. and Cooper, B.

Promoters for regulation of plant gene expression

Patent: WO 0198480-A 156 27-DEC-2001;

Syngenta Participations AG (CH)

FEATURES

source

1. .1217

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

ORIGIN

Query Match 82.0%; Score 1184; DB 6; Length 1217;

Best Local Similarity 99.3%; Pred. No. 2.7e-245;

Matches 1211; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

Qy 164 CAACTATTTTATGATGTCAGAGAGTCAGCATATGATTAATTCGATTCAGAAATCGTTTGAC 223

Db 1 CAACTATTTTATGATGTCAGAGAGTCAGCATATGATTAATTCGATTCAGAAATCGTTTGAC 60

Qy 224 GAGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283

Db 61 GAGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120

Qy 284 ATGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343

Db 121 ATGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180

Qy 344 TCAGGCTCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 403

Db 181 TCAGGCTCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240

Qy 404 TGTATGAATCTAAATGAACAGCCAAACAGCAGGAGCTAAACGTTGCTCGATGACT 463

Db 241 TGTATGAATCTAAATGAACAGCCAAACAGCAGGAGCTAAACGTTGCTCGATGACT 300

Qy 464 CGGTTTAAAGTTAAACCACTAAACCACTAAACCACTAAACCACTAAACCACTAAACCACT 523

Db 301 CGGTTTAAAGTTAAACCACTAAACCACTAAACCACTAAACCACTAAACCACTAAACCACT 360

Qy 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAGGGGTGAGATTAATAGTT 583

Db 361 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAGGGGTGAGATTAATAGTT 420

Qy 584 TAAAAATTAGTTAAACAGAGGGAAGAA- GCTGCTGACAGCGAGGTACAGTTATCTTTACC 642

Db 421 TAAAAATTAGTTAAACAGAGGGAAGAAAGGCTGCTGACAGCGAGGTACAGTTATCTTTACC 480

Qy 643 TGTGTCGAAATGATTCGTGCTGCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAA 702

Db 481 TGTGTCGAAATGATTCGTGCTGCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAA 540

Qy 703 AGTTGTAAGAGATAAACCCGCCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG 762

Db 541 AGTTGTAAGAGATAAACCCGCCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG 600

Qy 763 TCTCGTTGTCCTCCTCCTCCTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAG 822

Db 601 TCTCGTTGTCCTCCTCCTCCTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAG 660

Qy 823 AACAGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTT 882

Db 661 AACAGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGAGATTCATTTCTCGTTT 720

Qy 883 TGAATCTTCTCCTCAATCTCATCTTCTTCTCCGCTCTTCTTCCAGGTAATAGGAACTT 942

Db 721 TGAATCTTCTCCTCAATCTCATCTTCTTCTCCGCTCTTCTTCCAGGTAATAGGAACTT 777

Qy 943 TCTGGATCTACTTTATTTGCTGGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTGG 1002

Db 778 TCTGGATCTACTTTATTTGCTGGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTGG 837

Qy 1003 AATTCTGTTTAAATTTGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGAT 1062

Db 838 AATTCTGTTTAAATTTGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGAT 897

Qy 1063 CTAAGTTGACCGATCAGTTAGCTCGATATAGCTACCCAGAAATTTGGCTTGGATCTGG 1122

Db 898 CTAAGTTGACCGATCAGTTAGCTCGATATAGCTACCCAGAAATTTGGCTTGGATCTGG 957

Qy 1123 AGAGATCCATGTTTATGTTTACCTGGGAATGATTTGTATATGTGAATGAAATCTGAAT 1182

Db 958 AGAGATCCATGTTTATGTTTACCTGGGAATGATTTGTATATGTGAATGAAATCTGAAT 1017

Qy 1183 GTTGAAGTTAGATTTGAATCTGAACCTGCAATGTTAGATTTGAATCTGAACCTGTTAA 1242

Db 1018 GTTGAAGTTAGATTTGAATCTGAACCTGCAATGTTAGATTTGAATCTGAACCTGTTAA 1077

Qy 1243 GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGTTACGTTG 1302

Db 1078 GTTAGATGAAGTTTGTGTATAGATTTCTCGAACTTTAGGATTTGTAGTGTGTTACGTTG 1137

Qy 1303 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAATCTTTTGTG 1362

Db 1138 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAATCTTTTGTG 1197

Qy 1363 TGTGTTGAGCTCATAAAGAAAGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1422

Db 1198 TGTGTTGAGCTCATAAAGAAAGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1257

Qy 1423 TGGTACTGGAATGGT 1437

Db 1258 TGGTACCGGTATGGT 1272

Db	721	TGAATCTTCTCAATCTCA--TCTTCTTCGCGCTCTTTCTTTCCAAAGTAATAGGAACTT	777
Qy	943	TCGTGGATCTACTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCCCTTGAGATCTGG	1002
Db	778	TCGTGGATCTACTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCCCTTGAGATCTGG	837
Qy	1003	AATTCGTTTAAATTTGGATCTGTGAACTCCACTTAATCTTTTGGTTTACTAGAAATCGAT	1062
Db	838	AATTCGTTTAAATTTGGATCTGTGAACTCCACTTAATCTTTTGGTTTACTAGAAATCGAT	897
Qy	1063	CTAAGTTGACCGATCAGTTAGCTGATATAGCTACAGAAATTTGGCTTGACCTTTGATGG	1122
Db	898	CTAAGTTGACCGATCAGTTAGCTGATATAGCTACAGAAATTTGGCTTGACCTTTGATGG	957
Qy	1123	AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1182
Db	958	AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1017
Qy	1183	GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTTGAATTTGAATCTGAACACTGTTTAA	1242
Db	1018	GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTTGAATTTGAATCTGAACACTGTTTAA	1077
Qy	1243	GTTAGATGAAGTTTGTGTATAGATTTCTTGAACACTTGAATTTGTAGTGTCTGACGTTG	1302
Db	1078	GTTAGATGAAGTTTGTGTATAGATTTCTTGAACACTTGAATTTGTAGTGTCTGACGTTG	1137
Qy	1303	AACAGAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1362
Db	1138	AACAGAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1197
Qy	1363	TGTTTGCAGCTCATAAAAA 1382	
Db	1198	TGTTTGCAGCTCATAAAAA 1217	
RESULT 14			
AR236191			
LOCUS	AR236191	1219 bp	DNA linear PAT 20-DEC-2002
DEFINITION	Sequence 9 from patent US 6462258.		
ACCESSION	AR236191		
VERSION	AR236191.1 GI:27280000		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1219)		
AUTHORS	Fincher, K.L. and Wilkinson, J.Q.		
TITLE	Plant expression constructs		
JOURNAL	Patent: US 6462258-A 9 08-OCT-2002;		
FEATURES	Location/Qualifiers		
source	1..1219		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	81.4%; Score 1174.8; DB 6; Length 1219;		
Best Local Similarity	98.7%; Pred. No. 2.6e-243;		
Matches 1206; Conservative	0; Mismatches 12; Indels 4; Gaps 2;		
Qy	164	CAACTATTTTATGTATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTGGAC	223
Db	1	CAACTATTTTATGTATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTGGAC	60
Qy	224	GAGTTCGGATGTAGTAGTACCATTTAATTTAATGTACATACATACTAATCGTGAATAGTATG	283
Db	61	GAGTTCGGATGTAGTAGTACCATTTAATTTAATGTACATACATACTAATCGTGAATAGTATG	120
Qy	284	ATGAACATTTGTATCTTATTTGTATAAATCCATAAACACATCATGAAAGACACTTTCTT	343
Db	121	ATGAACATTTGTATCTTATTTGTATAAATCCATAAACACATCATGAAAGACACTTTCTT	180
Qy	344	TCAGGGTCTGAAATTAATATGATACAAATTTCTAATAGAAAAAGAAATTAATACGTTGAAT	403

RESULT 15  
AR438153

Db	181	TCACGGTCTGAATTAATTAATGATACAAATTTCTAATAGAAAAACGAA'TAAATTAACGTTGAAT	240
Qy	404	TGTATGAATCTTAATTTGAACAGCCAAACACAGCAGGAGCTAAACGTTGCCTGGATTGACT	463
Db	241	TGTATGAATCTTAATTTGAACAGCCAAACACAGCAGGAGCTAAACGTTGCCTGGATTGACT	300
Qy	464	CGTTTAAAGTTAAACCACTAAAAAAACGGAGCTGTCTATGTAAACACGCGGATCAGCAGGTC	523
Db	301	CGTTTAAAGTTAAACCACTAAAAAAACGGAGCTGTCTATGTAAACACGCGGATCAGCAGGTC	360
Qy	524	ACAGTCATGAAGCCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATAGTT	583
Db	361	ACAGTCATGAAGCCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATAGTT	420
Qy	584	TAAAAATTTAGTTTAAACACAGGGGAAA--GCTGTCTCAGCAGCCAGGTCAGGTTATCTTTACC	642
Db	421	TAAAAATTTAGTTTAAACACAGGGGAAAAGGCTGTCTCAGCAGCCAGGTCAGGTTATCTTTACC	480
Qy	643	TGTGTCGAAAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGGCCGAAATAA	702
Db	481	TGTGTCGAAAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGGCCGAAATAA	540
Qy	703	AGTTGTAAAGAGATAAAACCCGCTATATAAATTCATAATTTTCTCCCGCTTTGAAATG	762
Db	541	AGTTGTAAAGAGATAAAACCCGCTATATAAATTCATAATTTTCTCTCCGCTTTGAAATG	600
Qy	763	TCTCGTTGCTCCTCCTCAGCTTTCATCAGCGGTTTGAATCTCCGGGAGCTTGACAGAGA	822
Db	601	TCTCGTTGCTCCTCCTCAGCTTTCATCAGCGGTTTGAATCTCCGGGAGCTTGACAGAGA	660
Qy	823	AACAAGGAAGAGACTAAGAGAGAAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT	882
Db	661	AACAAGGAAGAGACTAAGAGAGAAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT	720
Qy	883	TGAATCTTCTCAATCTCATCTTCTTCTTCCGCTCTTCTTCCAAAGGTAATAGGAATCT	942
Db	721	TGAATCTTCTCAATCTCA--TCTTCTTCCGCTCTTCTTCCAAAGGTAATAGGAATCT	777
Qy	943	TCTGGATCTACATTTATTTGCTGGATCTCGATCTGTTTTCTCAATTTCTCTCAGATCTGG	1002
Db	778	TCTGGATCTACATTTATTTGCTGGATCTCGATCTGTTTTCTCAATTTCTCTCAGATCTGG	837
Qy	1003	AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAATCGAT	1062
Db	838	AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAATCGAT	897
Qy	1063	CTAAGTTGACCGATCAGTTAGCTCGATTTAGCTACGAGAAATTTGGCTTGACCTTGATGG	1122
Db	898	CTAAGTTGACCGATCAGTTAGCTCGATTTAGCTACGAGAAATTTGGCTTGACCTTGATGG	957
Qy	1123	AGAGATCCATGTTCTATGTTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1182
Db	958	AGAGATCCATGTTCTATGTTTACCTGGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1017
Qy	1183	GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTTTAA	1242
Db	1018	GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGTTTAA	1077
Qy	1243	GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTTCGTACGTTG	1302
Db	1078	GTTAGATGAAGTTTGTGTATAGATTTCTCGAAACCTTTAGGATTTGTAGTTCGTACGTTG	1137
Qy	1303	AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGATTTGAACTCTTTTGTG	1362
Db	1138	AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGATTTGAACTCTTTTGTG	1197
Qy	1363	TGTTTGCAGCTCATAAAAAATG 1384	
Db	1198	TGTTTGCAGCAGACTCACCATG 1219	

LOCUS	AR438153	1219 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	Sequence 9 from patent US 6660911.				
ACCESSION	AR438153				
VERSION	AR438153.1	GI:40204655			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1219)				
AUTHORS	Fincher,K.L., Flasinaki,S. and Wilkinson,J.Q.				
TITLE	Plant expression constructs				
JOURNAL	Patent: US 6660911-A 9 09-DEC-2003;				
FEATURES	Location/Qualifiers				
source	1..1219				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
Query Match	81.4%; Score 1174.8; DB 6; Length 1219;				
Best Local Similarity	98.7%; Pred. No. 2.6e-243;				
Matches 1206; Conservative	0; Mismatches 12; Indels 4; Gaps 2;				
QY	164	CAACTATTTTATGATGCAAGAGTCAGCATATGTATATATGATTTCAGAAATCGTTTTCGAC	223		
Db	1	CAACTATTTTATGATGCAAGAGTCAGCATATGTATATATGATTTCAGAAATCGTTTTCGAC	60		
QY	224	GAGTTCCGATGATGATGAGCCATTATTTAAATGTACATACTAATCGTGAATAGTGATATG	283		
Db	61	GAGTTCCGATGATGATGAGCCATTATTTAAATGTACATACTAATCGTGAATAGTGATATG	120		
QY	284	ATGAACATTTGATCTTATTTGTATATAATATCCATAAACACATCATCAATGAAAGACATTTCTT	343		
Db	121	ATGAACATTTGATCTTATTTGTATATAATATCCATAAACACATCATCAATGAAAGACATTTCTT	180		
QY	344	TCAGGCTCTGAATTAATTAATGATACAAATCTTAATAGAAACGAATTAATTAATGCTTCAAT	403		
Db	181	TCAGGCTCTGAATTAATTAATGATACAAATCTTAATAGAAACGAATTAATTAATGCTTCAAT	240		
QY	404	TGTATGAATCTAAATTTGAACCAAGCCAAACGACGAGGACTAAACGTTGCTCGATTTGACT	463		
Db	241	TGTATGAATCTAAATTTGAACCAAGCCAAACGACGAGGACTAAACGTTGCTCGATTTGACT	300		
QY	464	CGGTTTAAAGTTAAACCACTAAAAAACCGAGCTGTCTGTAAACACGGCGGATCCAGCAGGTC	523		
Db	301	CGGTTTAAAGTTAAACCACTAAAAAACCGAGCTGTCTGTAAACACGGCGGATCCAGCAGGTC	360		
QY	524	ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATAGTT	583		
Db	361	ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATAGTT	420		
QY	584	TAAAAATTAGTTTAAACGAGGGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTTACC	642		
Db	421	TAAAAATTAGTTTAAACGAGGGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTTACC	480		
QY	643	TGTGTCGAAATGATTCGTGTCGTGATTTTAAATTTATTTTGTGAAAGCCGAAAAATAA	702		
Db	481	TGTGTCGAAATGATTCGTGTCGTGATTTTAAATTTATTTTGTGAAAGCCGAAAAATAA	540		
QY	703	AGTTGCTAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG	762		
Db	541	AGTTGCTAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG	600		
QY	763	TCTCGTTGTCCTCCTCACCTTTTCATCAGCCGTTTTCGAAATCTCCGGGACCTTGACAGAGAAG	822		
Db	601	TCTCGTTGTCCTCCTCACCTTTTCATCAGCCGTTTTCGAAATCTCCGGGACCTTGACAGAGAAG	660		
QY	823	AACAAGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT	882		
Db	661	AACAAGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT	720		
QY	883	TGAATCTTCTCAATCTCATCTTCTTCTTCCGCTCTTCTTCCAGGTAATAGGAACCTT	942		
Db	721	TGAATCTTCTCAATCTCA---TCTTCTTCCGCTCTTCTTCCAGGTAATAGGAACCTT	777		

QY	943	TCTGGATCTACTTTTATTTTGGTGGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG	1002		
Db	778	TCTGGATCTACTTTTATTTTGGTGGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG	837		
QY	1003	AATTCGTTTTAAATTTGGGATCTGTGAAACCTCCACCTAAATCTTTTGGTTTTTACTAGAAATCGAT	1062		
Db	838	AATTCGTTTTAAATTTGGGATCTGTGAAACCTCCACCTAAATCTTTTGGTTTTTACTAGAAATCGAT	897		
QY	1063	CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTTACCAAGAAATTTGGCTTGACCTTGAATGG	1122		
Db	898	CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTTACCAAGAAATTTGGCTTGACCTTGAATGG	957		
QY	1123	AGAGATCCATGTTCAATGTTACCTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1182		
Db	958	AGAGATCCATGTTCAATGTTACCTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAACT	1017		
QY	1183	GTTGAAGTTAGATTCGAATCTGAACACCTGTCATGTTAGATTTGAATCTGAAACCTGTTAA	1242		
Db	1018	GTTGAAGTTAGATTCGAATCTGAACACCTGTCATGTTAGATTTGAATCTGAAACCTGTTAA	1077		
QY	1243	GTTAGATGAAGTTTGTGTATAGATTTCTTCGAAACCTTAGGATTTGTAGTCTCGTACGTTG	1302		
Db	1078	GTTAGATGAAGTTTGTGTATAGATTTCTTCGAAACCTTAGGATTTGTAGTCTCGTACGTTG	1137		
QY	1303	AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGACTGTATTTGAACTCTTTTGTG	1362		
Db	1138	AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGACTGTATTTGAACTCTTTTGTG	1197		
QY	1363	TGTTTGCAGCTCATAAAAAATG	1384		
Db	1198	TGTTTGCAGCAGACTCACCATG	1219		
Search completed: May 27, 2005, 22:19:52					
Job time : 7829 secs					

Search completed: May 27, 2005, 22:19:52  
Job time : 7829 secs



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 12:15:42 ; Search time 848 Seconds  
(without alignments)  
10080.314 Million cell updates/sec

Title: US-09-868-744B-1  
Perfect score: 1444  
Sequence: 1 attatgatctcaaatatcatt.....gtactggaatgtagtatcc 1444

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1:	Geneseqn1980s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004as:*
13:	Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	99.9	1444	3	AA61367 Arabidops
2	1341.8	92.9	6385	6	ABK89341 Plasmid p
3	1341.4	92.9	4526	6	ABK89340 Plasmid p
4	1306.2	90.5	11127	6	AA036967 Arabidops
5	1300.8	90.1	14184	12	ADL27876 E coli co
6	1298.4	89.9	3450	6	ABK89343 Plasmid p
7	1298.4	89.9	7474	12	ADN61599
8	1298.4	89.9	9356	12	ADN61613
9	1297.4	89.8	3408	6	ABK89342 Plasmid p
10	1297	89.8	2857	6	ABK89344
11	1174.8	81.4	1219	4	AA009791 Arabidops
12	1174	81.3	1742	4	AA009812 Chimeric
13	1114.6	77.2	12304	8	ABV75876 Luciferas
14	1104.6	76.5	1259	6	ABK52080 Modified
15	1078.8	74.7	15676	3	AA001288 Arabidops
16	1078.8	74.7	17111	3	AA001289 Arabidops
17	1078.8	74.7	17116	3	AA001290 Arabidops
18	1053	72.9	1202	6	ABK52077 Modified
19	1001.8	69.4	1342	6	ABK52081 Modified
20	950.2	65.8	1285	6	ABK52078 Modified

21	784.4	54.3	910	6	ABK52079 Modified
22	732.8	50.7	853	6	ABK52076 Modified
23	707.2	49.0	8573	12	ADO47655 Control o
24	579	40.1	1228	6	ABK53111 Transgene
25	475.8	33.0	11461	4	AA002175 Plasmid p
26	474.8	32.9	12766	4	AA002174 Plasmid p
27	163	11.3	1486	3	AA033486 Arabidops
28	116.4	8.1	1228	6	ABK53111 Transgene
29	110	7.6	573	10	ABK56906 Arabidops
30	94	6.5	94	6	ABK52089 Synthetic
31	84.4	5.8	97	6	ABK52087 Synthetic
32	79	5.5	79	6	ABK52084 Synthetic
33	77.4	5.4	79	6	ABK52092 Synthetic
34	76	5.3	2000	6	ABZ15497 Arabidops
35	75.4	5.2	1271	4	AA009792 Arabidops
36	73.8	5.1	1800	4	AA009811 Chimeric
37	65.4	4.5	67	6	ABK52085 Synthetic
38	65.4	4.5	74	6	ABK52082 Synthetic
39	63.4	4.4	77	6	ABK52083 Synthetic
40	59.2	4.1	1975	13	ADR73222 Thale cre
41	56.2	3.9	436	12	ADP93490 Cotton ex
42	56.2	3.9	2000	8	ADA71938 Rice gene
43	52.2	3.6	633	13	ACN52368 Cotton an
44	51.6	3.6	580	13	ADR63253 Cotton cd
45	51.2	3.5	1134	3	AAC45521 Arabidops

ALIGNMENTS

RESULT 1  
AAA61367  
ID AAA61367 standard; DNA; 1444 BP.  
XX  
AC AAA61367;  
XX  
DT 06-NOV-2000 (first entry)  
XX  
DE Arabidopsis thaliana actin promoter region.  
XX  
KW Actin 2; ACT2; promoter; oxalate oxidase; OXOX; insecticidal; fungicidal;  
KW antiviral; stability; db.  
XX  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers  
FT primer\_bind complement(1..21)  
FT primer\_bind /tag= a  
FT primer\_bind 1360..1379 /tag= b  
FT misc\_signal 1382..1384 /tag= c  
FT /note= "Translation start site for actin 2"  
XX  
XX WO200037661-A1.  
XX  
XX 29-JUN-2000.  
XX  
XX 16-DEC-1999; 99WO-GB004317.  
XX  
XX 21-DEC-1998; 98GB-00028201.  
XX  
XX (ADVA-) ADVANTA TECHNOLOGY LTD.  
XX  
XX Van Dun CMP, Schepers FMA, Pertijs JH;  
XX WPI; 2000-442681/38.  
XX  
XX Producing recombinant Compositae with increased transformation stability  
XX comprises linking the DNA construct to the ACT2 gene promoter.  
XX  
XX Claim 2; Fig 3; 22pp; English.

CC The present sequence is the promoter region and part of the coding  
CC sequence of the Arabidopsis gene, actin 2 (ACT2). The promoter may be  
CC used in a heterologous DNA construct to drive expression of RNA. This is  
CC useful for the production of recombinant plants containing genes with  
CC insecticidal, fungicidal or antiviral activity. The advantage of using  
CC the ACT2 gene promoter is that it increases the stability of the  
CC integrated DNA. In the present invention the ACT2 promoter was used to  
CC drive expression of the wheat oxalate oxidase gene  
XX  
SQ Sequence 1444 BP; 451 A; 241 C; 272 G; 478 T; 0 U; 2 Other;  
Query Match 99.9%; Score 1442; DB 3; Length 1444;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATTATGATCTCAATATACATGATACATATCTCATCTAGATCTAGGTATATATGTAAG 60  
DB 1 ATTATGATCTCAATATACATGATACATATCTCATCTAGATCTAGGTATATATGTAAG 60  
QY 61 AAAGTTTTCACGAATATGNNACGACAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120  
DB 61 AAAGTTTTCACGAATATGNNACGACAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120  
QY 121 TCAACATTATACTATACCAACATTTAGTTAGCAAAATTTAAACAACATATTTTATGTAAT 180  
DB 121 TCAACATTATACTATACCAACATTTAGTTAGCAAAATTTAAACAACATATTTTATGTAAT 180  
QY 181 GCAAGATCAGCATATGTAATTAATGTAATTCAGAAATCGTTTGACGATTCGGATGTAAG 240  
DB 181 GCAAGATCAGCATATGTAATTAATGTAATTCAGAAATCGTTTGACGATTCGGATGTAAG 240  
QY 241 TAGCCATTATTTAAATGTAATTAATGTAATTCAGAAATCGTTTGACGATTCGGATGTAAG 300  
DB 241 TAGCCATTATTTAAATGTAATTAATGTAATTCAGAAATCGTTTGACGATTCGGATGTAAG 300  
QY 301 ATTGTATAAATATCCATAAACAATCATGTAAGAAACATCTTTTTCAGGCTCTGAATTAAT 360  
DB 301 ATTGTATAAATATCCATAAACAATCATGTAAGAAACATCTTTTTCAGGCTCTGAATTAAT 360  
QY 361 TATGATACAAATTTCTAATAGAAAACGAATTAATTAATGTAATTCAGAAATCGTTTGAC 420  
DB 361 TATGATACAAATTTCTAATAGAAAACGAATTAATTAATGTAATTCAGAAATCGTTTGAC 420  
QY 421 AACAGCCAAACACGACGAGGACTAACGTTGCTGATGATGATGATGATGATGATGATGAT 480  
DB 421 AACAGCCAAACACGACGAGGACTAACGTTGCTGATGATGATGATGATGATGATGATGAT 480  
QY 481 TAAAAAACCAGAGCTGTCTATGTAACACGCGGATCGAGCAGGTACAGTCAATGAAGCCATC 540  
DB 481 TAAAAAACCAGAGCTGTCTATGTAACACGCGGATCGAGCAGGTACAGTCAATGAAGCCATC 540  
QY 541 AAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTAATTAATTAATTAATTAATTA 600  
DB 541 AAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTAATTAATTAATTAATTAATTA 600  
QY 601 GAGGGAAGAGCTGTCTGACAGCAGGTACGTTATCTTTACCTGTTGTCGAAATGATTCG 660  
DB 601 GAGGGAAGAGCTGTCTGACAGCAGGTACGTTATCTTTACCTGTTGTCGAAATGATTCG 660  
QY 661 TGTCTGTCGATTTAATTAATTTTGAAGGCCGAAAAATAAAGTTGTAAGAGATAAACCC 720  
DB 661 TGTCTGTCGATTTAATTAATTTTGAAGGCCGAAAAATAAAGTTGTAAGAGATAAACCC 720  
QY 721 CGCCTATATAAATTCATATAATTTTCTCCCGTTTGAATTTGTCGTTGTCCTCTCTAC 780  
DB 721 CGCCTATATAAATTCATATAATTTTCTCCCGTTTGAATTTGTCGTTGTCCTCTCTAC 780  
QY 781 TTTTCATCAGCGTTTGAATTTCTCCGCGACTTGACAGAGAAGAAACAGGAAGAGACTAA 840  
DB 781 TTTTCATCAGCGTTTGAATTTCTCCGCGACTTGACAGAGAAGAAACAGGAAGAGACTAA 840  
QY 841 GAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTTGAATTTCTCTCAATCTC 900  
DB 841 GAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTTGAATTTCTCTCAATCTC 900

DB 841 GAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTTGAATTTCTCTCAATCTC 900  
QY 901 ATCTCTCTCTCCGCTCTTTCTTTCCCAAGGTAATAGGAACCTTTCTGGAATCTACTTTATTT 960  
DB 901 ATCTCTCTCTCCGCTCTTTCTTTCCCAAGGTAATAGGAACCTTTCTGGAATCTACTTTATTT 960  
QY 961 GCTGATCTCGATCTGTTGTTTCTCAATTTCTTCTGAGATCTCGAATTCGTTTAAATTTGGAT 1020  
DB 961 GCTGATCTCGATCTGTTGTTTCTCAATTTCTTCTGAGATCTCGAATTCGTTTAAATTTGGAT 1020  
QY 1021 CTGTGAACCTCACATAAATCTTTTGGTTTCTAGATCTAGATCTAGATCTAGATCTAGATCTAG 1080  
DB 1021 CTGTGAACCTCACATAAATCTTTTGGTTTCTAGATCTAGATCTAGATCTAGATCTAGATCTAG 1080  
QY 1081 TAGCTCGAATATAGCTACCAAGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1140  
DB 1081 TAGCTCGAATATAGCTACCAAGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1140  
QY 1141 TACCTGGGAATGATTTGTATATGTAATTTGAAATCTGAACTCTGAACTCTGAACTCTGAACTCT 1200  
DB 1141 TACCTGGGAATGATTTGTATATGTAATTTGAAATCTGAACTCTGAACTCTGAACTCTGAACTCT 1200  
QY 1201 CTGAACACTGTCAATGTTAGATTTGAATCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCT 1260  
DB 1201 CTGAACACTGTCAATGTTAGATTTGAATCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCT 1260  
QY 1261 ATAGATCTCTGAAACCTTTAGATTTGTAGTGTGCTGATGATGATGATGATGATGATGATGAT 1320  
DB 1261 ATAGATCTCTGAAACCTTTAGATTTGTAGTGTGCTGATGATGATGATGATGATGATGATGAT 1320  
QY 1321 ATTCATCAGGCTTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1380  
DB 1321 ATTCATCAGGCTTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1380  
QY 1381 AATGCTGAGGCTGACGATATTTCAACCAATCGTGTGTCGACCAATGCTGATCTGGAATCGTAGG 1440  
DB 1381 AATGCTGAGGCTGACGATATTTCAACCAATCGTGTGTCGACCAATGCTGATCTGGAATCGTAGG 1440  
QY 1441 ATCC 1444  
DB 1441 ATCC 1444  
RESULT 2  
ABK89341  
ID ABK89341 standard; DNA; 6385 BP.  
XX  
AC ABK89341;  
XX  
DT 21-OCT-2002 (first entry)  
XX  
DE Plasmid pACT2B-GUS DNA.  
XX  
KW Arsenate reductase; ArsC; cyclic; circular; ds; antimonate; zinc;  
KW cadmium; phytochelatin biosynthetic enzyme; arsenite; arsenite; cobalt;  
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;  
KW industrial waste; phytoremediation.  
XX  
OS Synthetic.  
XX  
PN WO200248335-A2.  
XX  
PD 20-JUN-2002.  
XX  
PF 13-DEC-2001; 2001WO-US048105.  
XX  
PR 13-DEC-2001; 2000US-0255001P.  
XX  
PR 22-JUN-2001; 2001US-0300525P.  
XX  
PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX  
PI Meagher RB, Li Y;  
XX

DR WPI; 2002-593507/62.

XX Recombinant DNA molecules for producing transgenic plants tolerant to

PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of

PT contaminated soil or water, encodes arsenate reductase coding sequence.

XX Disclosure; Fig 13B; 131pp; English.

CC The invention relates to a nucleic acid molecule comprising a portion

CC encoding an arsenate reductase coding sequence and a plant-expressible

CC transcription regulatory sequence, the coding sequence being operably

CC linked to the transcription regulatory sequence. The nucleic acid is

CC useful for producing a plant which is resistant to at least one metal ion

CC such as a divalent cadmium ion or antimonate, by introducing the nucleic

CC acid molecule into a plant cell or into plant tissue, selecting for the

CC presence of the nucleic acid molecule to produce a transgenic plant cell

CC or plant tissue and regenerating a plant from the transgenic plant cell

CC or plant tissue. The method further comprises introducing at least one

CC plant-expressible phytochelatin biosynthetic enzyme coding sequence into

CC the cell to produce an arsenate metal ion resistant plant. The arsenate

CC reductase coding sequence is expressed under the control of a plant

CC promoter which directs expression in the above ground plant part. The

CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,

CC antimony, arsenate and arsenite ions. The transgenic plant is useful for

CC bioremediation of arsenate and/or cadmium-contaminated environments,

CC including soil, sediments, mine tailings, water, industrial waste,

CC groundwater and air. The transgenics are also useful for revegetation of

CC soils contaminated with metal ions and for removing and sequestering

CC these ions from water, wastewater and aqueous environments. The plants

CC are also useful for phytoremediation of contaminated soil, sediment,

CC water and mine tailings. This sequence represents plasmid pACT2B-GUS DNA,

CC used in the scope of the invention

XX

SQ Sequence 6385 BP; 1684 A; 1447 C; 1518 G; 1733 T; 0 U; 3 Other;

Query Match 92.9%; Score 1341.8; DB 6; Length 6385;

Best Local Similarity 98.8%; Pred. No. 5.8e-314; Mismatches 12; Indels 4; Gaps 2;

Matches 1375; Conservative 0;

QY 1 ATTATGATCTCAATATACATTGTATGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAAG 60

DB 257 ATTATGATCTCAATATACATTGTATGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAAG 316

QY 61 AAAGTTTGGACCAATATGNNACGACAAATGCTACACTCGATGTAATTTGGTATCTCAAC 120

DB 317 AAAGTTTGGACCAATATGNNACGACAAATGCTAGACTCGATGTAATTTGGTATCTCAAC 376

QY 121 TCACATATATCTTATACCAACATAGTTAGCAAAATTTTAAACACTATTTTATGTAT 180

DB 377 TCACATATATCTTATACCAACATAGTTAGCAAAATTTTAAACACTATTTTATGTAT 436

QY 181 GCAAGAGTCAGCATATGTATATATTCATTCAGAAATCGTTTTCAGAGTTCGGATGTAGTAG 240

DB 437 GCAAGAGTCAGCATATGTATATATTCATTCAGAAATCGTTTTCAGAGTTCGGATGTAGTAG 496

QY 241 TAGCCATTTATTAATGTATACATACTAATCGTGAATAGTATGATGAAACATTTGATCTTT 300

DB 497 TAGCCATTTATTAATGTATACATACTAATCGTGAATAGTATGATGAAACATTTGATCTTT 556

QY 301 ATTGTATATAATTCATTAACACATCATGAAAGACACTTTCTTTCAGGGTCTGAATTAAT 360

DB 557 ATTGTATATAATTCATTAACACATCATGAAAGACACTTTCTTTCAGGGTCTGAATTAAT 616

QY 361 TATGATACAAATTTCTAATAGAAAACGAATTAATATACGTTTGAATTCGTATGAATCTAAATTG 420

DB 617 TATGATACAAATTTCTAATAGAAAACGAATTAATATACGTTTGAATTCGTATGAATCTAAATTG 676

QY 421 AACAGCCCAACCAACGACGAGGACTAACCGTTGCTCGATTTGACTCGGTTTAAAGTTAACCCAC 480

DB 677 AACAGCCCAACCAACGACGAGGACTAACCGTTGCTCGATTTGACTCGGTTTAAAGTTAACCCAC 736

QY 481 TAAAAAAGCGAGCTGTCTATGTAAACACCGCGGATCGAGCAGGTCACAGTCATGAAGCCATC 540

DB 737 TAAAAAAGCGAGCTGTCTATGTAAACACCGCGGATCGAGCAGGTCACAGTCATCAAGCCCATC 796

QY 541 AAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAATTAATTAATTAATTAATTAATTAAC 600

DB 797 AAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAATTAATTAATTAATTAATTAAC 856

QY 601 AAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGTGTCGAAATGATTC 659

DB 857 AAGGAAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGTGTCGAAATGATTC 916

QY 660 GTGTCTGTGCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAAAGTTGTAAGAGATAAAC 719

DB 917 GTGTCTGTGCTGATTTTAAATTTTAAATTTTAAAGGCGCGAAATAAAGTTGTAAGAGATAAAC 976

QY 720 CGCCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGTCCTCTCA 779

DB 977 CGCCCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTCGTTGTCCTCTCA 1036

QY 780 CTTTCATCAGCGGTTTGAATTTCTCCGGCAGCTTGTACAGAGAAACAAGAAAGAGACTA 839

DB 1037 CTTTCATCAGCGGTTTGAATTTCTCCGGCAGCTTGTACAGAGAAACAAGAAAGAGACTA 1096

QY 840 AGAGAGAAAGTAAGAGATAAATCCAGGAGATTCATTTCTCGTTTGAATTTCTCTCAATCT 899

DB 1097 AGAGAGAAAGTAAGAGATAAATCCAGGAGATTCATTTCTCCGTTTGAATTTCTCTCAATCT 1156

QY 900 CATCTTCTTTCTCCGCTCTTTCTTTTCCAAAGGTAATAGGAACCTTTCTGGATCTACTTTAT 959

DB 1157 CA---TCCTTCTCCGCTCTTTCTTTTCCAAAGGTAATAGGAACCTTTCTGGATCTACTTTAT 1213

QY 960 TGTGGAATCTGATCTGTTTCTCAATTTCTCTGAGATCTGGAATTCGTTTAAATTTGGA 1019

DB 1214 TGTGGAATCTGATCTGTTTCTCAATTTCTCTGAGATCTGGAATTCGTTTAAATTTGGA 1273

QY 1020 TCTGTGAACCTCCACTAATCTTTTGGTTTACTAGATCGATCTAAGTTGACCGATCAG 1079

DB 1274 TCTGTGAACCTCCACTAATCTTTTGGTTTACTAGATCGATCTAAGTTGACCGATCAG 1333

QY 1080 TTAGCTCGATTTATAGTACACAGAAATTTGGCTTTGACCTTGATGGAGAGATCCATGTTTCATG 1139

DB 1334 TTAGCTCGATTTATAGTACACAGAAATTTGGCTTTGACCTTGATGGAGAGATCCATGTTTCATG 1393

QY 1140 TTACTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAAATCTGTTTGAAGTTAGATGAA 1199

DB 1394 TTACTCGGAAATGATTTGTATATGTGAATTTGAAATCTGAAATCTGTTTGAAGTTAGATGAA 1453

QY 1200 TCTGAACTCTCAATGTAGATTTGAATTTGAATCTGAACTGTTTAAAGTTAGATGAAATTTG 1259

DB 1454 TCTGAACTCTCAATGTAGATTTGAATTTGAATCTGAACTGTTTAAAGTTAGATGAAATTTG 1513

QY 1260 TATAGATTTCTCGAAACCTTAGGATTTTGTAGTGTCTGATGTTGAACAGAAAGCTATTTCT 1319

DB 1514 TATAGATTTCTCGAAACCTTAGGATTTTGTAGTGTCTGATGTTGAACAGAAAGCTATTTCT 1573

QY 1320 GATTCAATCAGGGTTATTTGACTGTATGAACTCTTTTGTGTGTTTGTGAGCTCATAAA 1379

DB 1574 GATTCAATCAGGGTTATTTGACTGTATGAACTCTTTTGTGTGTTTGTGAGCTCATAAA 1633

QY 1380 AAATGGCTGAG 1390

DB 1634 GGATCCCGGG 1644

RESULT 3

ABK89340

ID ABK89340 standard; DNA; 4526 BP.

XX

AC ABK89340;

XX

DT 21-OCT-2002 (first entry)

XX

DE Plasmid pACT2B DNA.

XX

KW Arsenate reductase; ArsC; cyclic; circular; ds; antimonate; zinc;  
KW cadmium; phytochelatin biosynthetic enzyme; arsinic acid; arsenite; cobalt;  
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;  
KW industrial waste; phytoremediation.  
XX Synthetic.  
OS  
XX WO200248335-A2.  
XX  
XX 20-JUN-2002.  
XX  
XX 13-DEC-2001; 2001WO-US048105.  
XX  
XX 13-DEC-2000; 2000US-0255001P.  
PR 22-JUN-2001; 2001US-0300525P.  
XX  
XX (UYGE-) UNIV GEORGIA RES FOUND INC.  
PA  
XX Meagher RB, Li Y;  
PI  
XX WPI; 2002-583507/62.  
XX  
XX Recombinant DNA molecules for producing transgenic plants tolerant to  
PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of  
PT contaminated soil or water, encodes arsenate reductase coding sequence.  
XX  
XX Disclosure; Fig 12B; 131pp; English.  
XX  
XX The invention relates to a nucleic acid molecule comprising a portion  
CC encoding an arsenate reductase coding sequence and a plant-expressible  
CC transcription regulatory sequence, the coding sequence being operably  
CC linked to the transcription regulatory sequence. The nucleic acid is  
CC useful for producing a plant which is resistant to at least one metal ion  
CC such as a divalent cadmium ion or antimonate, by introducing the nucleic  
CC acid molecule into a plant cell or into plant tissue, selecting for the  
CC presence of the nucleic acid molecule to produce a transgenic plant cell  
CC or plant tissue and regenerating a plant from the transgenic plant cell  
CC or plant tissue. The method further comprises introducing at least one  
CC plant expressible phytochelatin biosynthetic enzyme coding sequence into  
CC the cell to produce an arsenate metal ion resistant plant. The arsenate  
CC reductase coding sequence is expressed under the control of a plant  
CC promoter which directs expression in the above ground plant part. The  
CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,  
CC antimony, arsenate and arsenite ions. The transgenic plant is useful for  
CC bioremediation of arsenate and/or cadmium-contaminated environments,  
CC including soil, sediments, mine tailings, water, industrial waste,  
CC groundwater and air. The transgenics are also useful for revegetation of  
CC soils contaminated with metal ions and for removing and sequestering  
CC these ions from water, wastewater and aqueous environments. The plants  
CC are also useful for phytoremediation of contaminated soil, sediment,  
CC water and mine tailings. This sequence represents plasmid pACT2B DNA,  
CC used in the scope of the invention  
XX  
SQ Sequence 4526 BP; 1203 A; 997 C; 1002 G; 1321 T; 0 U; 3 Other;

Query Match 92.9%; Score 1341.4; DB 6; Length 4526;  
Best Local Similarity 98.9%; Pred. No. 5.8e-314;  
Matches 13/4; Conservative 0; Mismatches 11; Indels 4; Gaps 2;

QY 1 ATTATGATCTCAATATGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAG 60  
DB  
DB 257 ATTATGATCTCAATATGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAG 316  
QY 61 AAAGTTTTCAGCAATATGNNACGAAATGGCTACACTCGATGTAAATGGTATCTCAAC 120  
DB 317 AAAGTTTTCAGCAATATGNNACGAAATGGCTAGACTCGATGTAAATGGTATCTCAAC 376  
QY 121 TCACATTATCTTATACCAACATTTAGTACCAAAATTTAAACACTATTTTATGTAAT 180  
DB 377 TCACATTATCTTATACCAACATTTAGTACCAAAATTTAAACACTATTTTATGTAAT 436  
QY 181 GCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTTGACGAGTTCGAGTAGTAG 240

DB 437 GCAGAGTCAGCATATGTATTAATTTGATTCAGAAATCGTTTTGACGAGTTCGAGTAGTAG 496  
QY 241 TAGCCATTATTTAATGTATACATACATAATCGTGAATAGTATGATGAACAACTGTTATCTT 300  
DB 497 TAGCCATTATTTAATGTATACATACATAATCGTGAATAGTATGATGAACAACTGTTATCTT 556  
QY 301 ATTGTATAAATATCCATAAACAACATCATGAAAGACACATTTCTTTCCAGGCTCTGAATTAAT 360  
DB 557 ATTGTATAAATATCCATAAACAACATCATGAAAGACACATTTCTTTCCAGGCTCTGAATTAAT 616  
QY 361 TATGATACAAATTTAATAGAAAAAAGAAATTAATTTAGTTGTAATTTGTAATTAATTTG 420  
DB 617 TATGATACAAATTTAATAGAAAAAAGAAATTAATTTAGTTGTAATTTGTAATTTG 676  
QY 421 AACAGCCCAACACGACGAGGACTAAACGTTGCTCGATTTGACTCGGTTTAAAGTTAAACAC 480  
DB 677 AACAGCCCAACACGACGAGGACTAAACGTTGCTCGATTTGACTCGGTTTAAAGTTAAACAC 736  
QY 481 TAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCGAGCAGGTTCACAGTCATGAAGCCATC 540  
DB 737 TAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCGAGCAGGTTCACAGTCATGAAGCCATC 796  
QY 541 AAAGCAAAAGAACTAATCCAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAAAC 600  
DB 797 AAAGCAAAAGAACTAATCCAGGGGTGAGATGAATTAATTTAGTTTAAAAATTTAGTTAAAC 856  
QY 601 GAGGAAAAA-GCTGTCTGCACAGCCAGGTTCAGTTATCTTTTACCTGTGTGCGAAATGATTC 659  
DB 857 GAGGAAAAAAGGCTGTCTGCACAGCCAGGTTCAGTTATCTTTTACCTGTGTGCGAAATGATTC 916  
QY 660 GTGCTGTCTGATTTTAAATTTTAAATTTTAAAGGCGGAAAAAATAAGTTTGAAGAGATAAC 719  
DB 917 GTGCTGTCTGATTTTAAATTTTAAATTTTAAAGGCGGAAAAAATAAGTTTGAAGAGATAAC 976  
QY 720 CGCCTATATAAATTCATATATTTTCTCCCGCTTTGAAATTTGCTGTTGCTCTCTCA 779  
DB 977 CGCCTATATAAATTCATATATTTTCTCTCGCTTTGAAATTTGCTGTTGCTCTCTCA 1036  
QY 780 CTTTCATCAGCGTTTTCGAATCTCCGCGACTTGACAGAGAAACAAGGAAGAAGACTA 839  
DB 1037 CTTTCATCAGCGTTTTCGAATCTCCGCGACTTGACAGAGAAACAAGGAAGAAGACTA 1096  
QY 840 AGAGAGAAAGTAAGAGATAATTCAGGAGATTCATTTCCGTTTGAATCTTCTCTCAATCT 899  
DB 1097 AGAGAGAAAGTAAGAGATAATTCAGGAGATTCATTTCCGTTTGAATCTTCTCTCAATCT 1156  
QY 900 CATCTCTCTTCCGCTCTTTCTTTCCAGGTAATAGGAACCTTTCTGGATCTACTTTATT 959  
DB 1157 CA--TCTTCTTCCGCTCTTTCTTTCCAGGTAATAGGAACCTTTCTGGATCTACTTTATT 1213  
QY 960 TGCTCGATCTCGATCTGTTTCTCAATTTCTCTGAGATCTGGAATTCGTTTAAATTTGGA 1019  
DB 1214 TGCTCGATCTCGATCTGTTTCTCAATTTCTCTGAGATCTGGAATTCGTTTAAATTTGGA 1273  
QY 1020 TCTGTGAACCTCCACTAAATCTTTTGGTTTCTAGATAATTCGATCTAAGTTGACCGATCAG 1079  
DB 1274 TCTGTGAACCTCCACTAAATCTTTTGGTTTCTAGATAATTCGATCTAAGTTGACCGATCAG 1333  
QY 1080 TTAGTCTGATTTATAGTACCAAGATTTGGCTTTGACCTTGAATGGAGAGATCCATGTTGAT 1139  
DB 1334 TTAGTCTGATTTATAGTACCAAGATTTGGCTTTGACCTTGAATGGAGAGATCCATGTTGAT 1393  
QY 1140 TTAGTCTGGAATGATTTGATATGTAATTTGAAATCTGAACTTTGAAGTTAGATTGAA 1199  
DB 1394 TTAGTCTGGAATGATTTGATATGTAATTTGAAATCTGAACTTTGAAGTTAGATTGAA 1453  
QY 1200 TCTGAACACTGTCTCAATTTAGATTTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTG 1259  
DB 1454 TCTGAACACTGTCTCAATTTAGATTTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTG 1513  
QY 1260 TATAGATTTCTGAAACCTTAGGATTTTGTAGTGTGTTGAGCTTTGAACAGAACTATTCT 1319  
DB 1514 TATAGATTTCTGAAACCTTTAGGATTTTGTAGTGTGTTGAGCTTTGAAACAGAACTATTCT 1573

```
QY 1320 GATTCAATCAGGGTTTATTGACTGTAATGAACTCTTTTGTGTGTTGCGAGTCATAAA 1379
Db |||||
QY 1574 GATTCAATCAGGGTTTATTGACTGTAATGAACTCTTTTGTGTGTTGCGAGTCATAAA 1633
Db |||||
QY 1380 AAATGGCTG 1388
Db |||||
Db 1634 GGATCCCG 1642

RESULT 4
AAD36967
ID AAD36967 standard; DNA; 11127 BP.
AC AAD36967;
XX
XX 29-AUG-2003 (revised)
DT 21-AUG-2002 (first entry)
XX
DE Arabidopsis actin-2 promoter-H04 toxin chimeric construct, pZU578.
KW Hybrid delta toxin H04; CryIAb toxin; CryIC toxin; transgenic plant;
KW transgenic; insect control; insect resistance; agricultural; insecticide;
KW actin-2 promoter; pZU578 construct; ds.
XX
OS Bacillus thuringiensis.
OS Arabidopsis sp.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT promoter 56..1475
FT /*tag= a
FT /note= "Actin-2 promoter"
FT misc_feature 1485..1491
FT /*tag= b
FT /note= "Encodes the toxin portion of H04 and the first 40
FT amino acids of CryIAb tail"
FT promoter 3859..5030
FT /*tag= c
FT /note= "Maize ubiquitin promoter"
FT misc_feature 5052..6271
FT /*tag= d
FT /note= "Phosphomannose isomerase (PMI) marker gene"
XX
XX WO200215701-A2.
XX
XX 28-FEB-2002.
XX
XX 23-AUG-2001; 2001WO-EP009751.
XX
XX 25-AUG-2000; 2000US-0227956P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Carozzi NB, Rabe SM, Miles PJ, Warren GW, De Haan PT;
XX
XX WPI; 2002-280838/32.
XX
XX Bacillus thuringiensis hybrid toxin H04 for controlling insects, e.g.
XX fall army worm or European cornborer, and for creating insect resistant
XX plants, comprises domains I and II of CryIAb and domain III of CryIC.
XX
XX Claim 25; Page 127-130; 130pp; English.
XX
XX The present invention relates to Bacillus thuringiensis hybrid delta
XX toxin H04 comprising a N-terminal toxin portion which contain domains I
XX and II from CryIAb toxin joined in the amino to carboxy direction to
XX domain III from CryIC toxin and a C-terminal tail region from a CryIAb
XX toxin. The B. thuringiensis toxins are useful in multiple insect control
XX strategies by inhibiting the ability of insect pests to survive, grow or
XX reproduce or of limiting insect-related damage or loss in crop plants.
XX Nucleotide sequences encoding the toxin are used to generate transgenic
XX plants having resistance to insects. The toxins may be applied in insect
```

```
CC infested areas or to prophylactically treat insect susceptible areas or
CC plants to confer protection or resistance against harmful insects. The
CC present DNA sequence is pZU578 chimeric construct comprising Arabidopsis
CC actin-2 promoter operatively linked to the B. thuringiensis H04 toxin
CC portion and the first 40 amino acids of CryIAb tail. (Updated on 29-AUG-
CC 2003 to standardise OS field)
XX
SQ Sequence 11127 BP; 2866 A; 2919 C; 2758 G; 2584 T; 0 U; 0 Other;
```

```
Query Match 90.5%; Score 1306.2; DB 6; Length 11127;
Best Local Similarity 98.8%; Pred. No. 2.6e-305;
Matches 1368; Conservative 0; Mismatches 10; Indels 7; Gaps 5;

QY 1 ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGTATTATTTATGTAAG 60
Db |||||
QY 93 ATTATGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGTATTATTTATGTAAG 152
Db |||||
QY 61 AAAGTTTGGAGCAATATGNNAGCAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
Db |||||
QY 153 AAAGTTTGGAGCAATATGNNAGCAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 212
Db |||||
QY 121 TCAACATTATATCTTATACCAAACTAGTTAG-CAAAATTTTAAACAACTA-TTTTATGT 178
Db |||||
QY 213 TCAACATTATATCTTATACCAAACTAGTTAG-CAAAATTTTAAACAACTA-TTTTATGT 272
Db |||||
QY 179 ATGCAAGAGTCAGCATATGTATAATTGATTCAGAAATCGTTTTCGAGTTCGGATGTAGT 238
Db |||||
QY 273 ATGCAAGAGTCAGCATATGTATAATTGATTCAGAAATCGTTTTCGAGTTCGGATGTAGT 332
Db |||||
QY 239 AGTAGCCATTTATTTAATGTACATATCTAATCGTGAATAGTG-ATATGATGAACATTTGAT 297
Db |||||
QY 333 AGTAGCCATTTATTTAATGTACATATCTAATCGTGAATAGTG-ATATGATGAACATTTGAT 392
Db |||||
QY 298 CTTATTGTATAAATATCCATAAACAACATCATGAAGACACTTTCTTTCAGGCTCTGAATT 357
Db |||||
QY 393 CTTATTGTATAAATATCCATAAACAACATCATGAAGACACTTTCTTTCAGGCTCTGAATT 452
Db |||||
QY 358 AATTATGATACAATTTCTAATAGAAAACGAATTAATTAACGTTTGAATTTGATGAATCTAA 417
Db |||||
QY 453 AATTATGACACAATTTCTAATAGAAAACGAATTAATTAACGTTTGAATTTGATGAATCTAA 512
Db |||||
QY 418 TTGAACAAAGCCAAACCAACGACGAGCACTAACCGTTGCTGGATGACTCGGTTAAGTTAAC 477
Db |||||
QY 513 TTGAACAAAGCCAAACCAACGACGAGCACTAACCGTTGCTGGATGACTCGGTTAAGTTAAC 572
Db |||||
QY 478 CACTAAAACAAACGAGGCTGTCTATGTAACACCGCGATCGAGCAGGTTCACAGTCATCAAGCC 537
Db |||||
QY 573 CACTAAAACAAACGAGGCTGTCTATGTAACACCGCGATCGAGCAGGTTCACAGTCATCAAGCC 632
Db |||||
QY 538 ATCAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTAAGTTTAAAAATTTAGTTAA 597
Db |||||
QY 633 ATCAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTAAGTTTAAAAATTTAGTTAA 692
Db |||||
QY 598 CACGAGGGAATAA-GCTGTCTGACAGCCAGGTTCAGTTATCTTTTACCTGTGGTCGAAATGA 656
Db |||||
QY 693 CACGAGGGAATAAAGGCTGTCTGACAGCCAGGTTCAGTTATCTTTTACCTGTGGTCGAAATGA 752
Db |||||
QY 657 TTCGTGCTGTGCGATTTTAAATTTTTCGAAAGGCGGAAATAAAGTGTGTAAGAGATA 716
Db |||||
QY 753 TTCGTGCTGTGCGATTTTAAATTTTTCGAAAGGCGGAAATAAAGTGTGTAAGAGATA 812
Db |||||
QY 717 AACCGCGCTATATAAAATTCATATATTTTCCTCCCGCTTTTGAATTTGCTCGTGTGCTCC 776
Db |||||
QY 813 AACCGCGCTATATAAAATTCATATATTTTCCTCCCGCTTTTGAATTTGCTCGTGTGCTCC 872
Db |||||
QY 777 TCACCTTTCATCAGCGGTTTGAATCTCGGGCAGCTTGACAGAGAAACAAGAGAAAGA 836
Db |||||
QY 873 TCACCTTTCATCAGCGGTTTGAATCTCGGGCAGCTTGACAGAGAAACAAGAGAAAGA 932
Db |||||
QY 837 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGGTTTGAATTTCTCTCAA 896
Db |||||
QY 933 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGGTTTGAATTTCTCTCAA 992
Db |||||
```

QY 897 TCTCATCTTCTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTCTCGATCTACTTTT 956  
|||||  
Db 993 TCTCA---TCTTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTCTCGATCTACTTT 1049  
  
QY 957 ATTGCTGAGCTCGATCTGTGTTTCTCAATTTCTTGGATCTCGAATTCGTTAAATTT 1016  
Db 1050 ATTGCTGAGCTCGATCTGTGTTTCTCAATTTCTTGGATCTCGAATTCGTTAAATTT 1109  
  
QY 1017 GATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGAACTGAATCGATCTAAGTTGACCGAT 1076  
Db 1110 GGATCTGTGAACCTCCACTAAATCTTTGGTTTACTAGAACTGAATCGATCTAAGTTGACCGAT 1169  
  
QY 1077 CAGTTAGCTCGATTTAGCTACAGAAATTGGCTTTGACCTTGCATGGAGAGATCCCATGTTTC 1136  
Db 1170 CAGTTAGCTCGATTTAGCTACAGAAATTGGCTTTGACCTTGCATGGAGAGATCCCATGTTTC 1229  
  
QY 1137 ATGTTACCTGGAAATGATTTGTATATGTGAATGAAATCTGAACTGTGTTGAAAGTTAGATT 1196  
Db 1230 ATGTTACCTGGAAATGATTTGTATATGTGAATGAAATCTGAACTGTGTTGAAAGTTAGATT 1289  
  
QY 1197 GAATCTGAACACTGTCMAATGTTAGATTGAACTGACACTGTTTAAAGTTAGATGAATTT 1256  
Db 1290 GAATCTGAACACTGTCMAATGTTAGATTGAACTGACACTGTTTAAAGTTAGATGAATTT 1349  
  
QY 1257 GTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGACGTTTGAACAGAAAGCTATT 1316  
Db 1350 GTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGACGTTTGAACAGAAAGCTATT 1409  
  
QY 1317 TCTGATTCATCAGGGTTTATTGACTGATGATTTGAACTCTTTTTTGTGTTTCAGCTCAT 1376  
Db 1410 TCTGATTCATCAGGGTTTATTGACTGATGATTTGAACTCTTTTTTGTGTTTCAGCTCAT 1469  
  
QY 1377 AAAAA 1381  
Db 1470 AAAAA 1474

RESULT 5

ADL27876  
ID ADL27876 standard; DNA; 14184 BP.  
XX  
AC ADL27876;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE E coli codA gene vector #3.  
XX  
ds; gene; vector; selection method; yield; pathogen resistance;  
KW nutritional quality; stress resistance; plant; transgenic.  
XX  
OS Escherichia coli.  
OS Synthetic.  
XX  
PN WO2004013333-A2.  
XX  
PD 12-FEB-2004.  
XX  
PF 18-JUL-2003; 2003WO-BP007877.  
XX  
PR 26-JUL-2002; 2002DE-01034287.  
XX  
PA (BADI ) BASF PLANT SCI GMBH.  
XX  
PI Kock M, Frank M, Badur R;  
XX  
DR WPI; 2004-157134/15.  
XX  
PT Method for transforming plant cells, useful for preparing transgenic  
PT plants for e.g. food or production of chemicals, with selection based on  
PT suppressing toxic effects of a marker protein.  
XX  
PS Disclosure; Page 189-194; 201pp; German.  
XX

CC The present invention relates to a method for preparing transformed plant  
CC cells or organisms, which comprises first transfecting a population of  
CC plant cells, already containing at least one marker protein that can  
CC exert a direct or indirect toxic effect, with at least one nucleic acid  
CC sequence together with at least one compound able to reduce expression,  
CC amount, activity and/or function of the marker protein. Transfected cells  
CC that contain the nucleic acid in their genome and, because of the  
CC activity of the compound have a growth advantage relative to non-  
CC transformed cells are selected, under conditions where the marker protein  
CC exerts its toxic effect on non-transformed cells. The method is  
CC especially used to produce transgenic plants, especially those having an  
CC advantageous phenotype, e.g. better nutritional quality for humans or  
CC animals, production of selected chemicals or pharmaceuticals, increased  
CC resistance to pathogens or environmental stress or higher yields. The  
CC present sequence is a coding sequence shown in the exemplification of the  
CC invention.

XX  
SQ Sequence 14184 BP; 3818 A; 3314 C; 3385 G; 3665 T; 0 U; 2 Other;  
Query Match 90.1%; Score 1300.8; DB 12; Length 14184;  
Best Local Similarity 98.1%; Pred. No. 5.6e-304;  
Matches 1379; Conservative 0; Mismatches 19; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAATPACATTCATCTAGATCTAGATCTAGTTATCATTTATGTAAG 60  
Db 9980 ATTATGATCTCAATPACATTCATCTAGATCTAGATCTAGTTATCATTTATGTAAG 10039  
  
QY 61 AAAGTTTTGACGATATGNNACGCAAAATGGCTACACTCGATGTAATGGTATCTCAAC 120  
Db 10040 AAAGTTTTGACGATATGNNACGCAAAATGGCTAGACTCGATGTAATGGTATCTCAAC 10099  
  
QY 121 TCAACATTTATCTTATACCAACATTTAGTTAG-CAAAAATTTAAACAACTA-TTTTTATGT 178  
Db 10100 TCAACATTTATCTTATACCAACATTTAGTTAGCAAAATTTAAACAACTATTTTTTATGT 10159  
  
QY 179 ATGCAAGAGTCAGCATATGTATAATTTGAATTCAGAACTGTTTGACAGTTCGGATGTAGT 238  
Db 10160 ATGCAAGAGTCAGCATATGTATAATTTGAATTCAGAACTGTTTGACAGTTCGGATGTAGT 10219  
  
QY 239 AGTAGCCATTTATTTATGTCATCTACTAATCTGATAGTGTG-ATATGATGCAACATTTGAT 297  
Db 10220 AGTAGCCATTTATTTATGTCATCTACTAATCTGATAGTGTGATGATGAAACATTTGAT 10279  
  
QY 298 CTTATTGTATAAATCCATAAACAACATCATGAAAGACACTTTCTTCAGGGTCTGAAAT 357  
Db 10280 CTTATTGTATAAATCCATAAACAACATCATGAAAGACACTTTCTTCAGGGTCTGAAAT 10339  
  
QY 358 AATTATGATACAATTTCTAATAGAAAAAGAAATTAATTAACGTTGAAATTTGATGAAATCTAA 417  
Db 10340 AATTATGATACAATTTCTAATAGAAAAAGAAATTAATTAACGTTGAAATTTGATGAAATCTAA 10399  
  
QY 418 TTGAACAAGCCCAACCCAGCAGGACTAAACGTTGCTCGATTGACTCCGTTTAAAGTTAAC 477  
Db 10400 TTGAACAAGCCCAACCCAGCAGGACTAAACGTTGCTCGATTGACTCCGTTTAAAGTTAAC 10459  
  
QY 478 CACTAAAAAAGCGAGCTGTATGTAAACAACGCGGATCGAGAGGTCACAGTCATCAAGCC 537  
Db 10460 CACTAAAAAAGCGAGCTGTATGTAAACAACGCGGATCGAGAGGTCACAGTCATCAAGCC 10519  
  
QY 538 ATCAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTAAGTTAA 597  
Db 10520 ATCAAAGCAAAAGAACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAAAATTAAGTTAA 10579  
  
QY 598 CACGAGGGGAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGGTCGAAATGA 656  
Db 10580 CACGAGGGGAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGGTCGAAATGA 10639  
  
QY 657 TTCGTGTCGTGCTGATTTTAAATTTATTTTGTAAAGCCGCAAAATAAAGTTGTAAGAGATA 716  
Db 10640 TTCGTGTCGTGCTGATTTTAAATTTATTTTGTAAAGCCGCAAAATAAAGTTGTAAGAGATA 10699  
  
QY 717 AACCCGCTTATATAAAATTCATATATTTTCTCCCGCTTTTGAATTTGCTGTTGTCCTCC 776  
|||||



Db 10700 AACCCGCTATATATAATTCATATATTTCTCTCCGCTTTGAAATGTCTCGTTGCTCTCC 10759  
QY 777 TCACCTTTTCATAGCCGCTTTTGAATCTCCGCGCATTTGACAGAGAAAGAAACAAGAAAGA 836  
Db 10760 TCACCTTTTCATAGCCGCTTTTGAATCTCCGCGCATTTGACAGAGAAAGAAACAAGAAAGA 10819  
QY 837 CTAAGAGAGAAAGTAAGAGATATCCAGAGAGATTCATCTCCGTTTGAATCTTCTCTCAA 896  
Db 10820 CTAAGAGAGAAAGTAAGAGATATCCAGAGAGATTCATCTCCGTTTGAATCTTCTCTCAA 10879  
QY 897 TCTCATCT 956  
Db 10880 TCTCA---TCT 10936  
QY 957 ATTGCTGAGATCGATCT 1016  
Db 10937 ATTGCTGAGATCGATCT 10996  
QY 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGATCGATCTAAAGTTGACCGAT 1076  
Db 10997 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGATCGATCTAAAGTTGACCGAT 11056  
QY 1077 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTTGACCTTCATCGAGAGATCCATGTTTC 1136  
Db 11057 CAGTTAGCTCGATATAGCTACAGAAATTTGGCTTTGACCTTCATCGAGAGATCCATGTTTC 11116  
QY 1137 ATGTTACCTGGGAATGATTTGTATATGTGAATGAAATCTGAACTGTTGAAATTTAGATT 1196  
Db 11117 ATGTTACCTGGGAATGATTTGTATATGTGAATGAAATCTGAACTGTTGAAATTTAGATT 11176  
QY 1197 GAATCTGAACACTGCAATGTAGATTTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255  
Db 11177 GAATCTGAACACTGCAATGTAGATTTGAATCTGAACACTGTTTAAAGTTAGATGAAGTT 11236  
QY 1256 TGTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGACGTTGACAGAGAGCTAT 1315  
Db 11237 TGTGTATAGATTTCTCGAAACCTTTAGGATTTGTAGTGTGACGTTGACAGAGAGCTAT 11296  
QY 1316 TTCTGATTCATFCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTGTGTTTGCAGCTCA 1375  
Db 11297 TTCTGATTCATFCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTGTGTTTGCAGCTCA 11356  
QY 1376 TAAAAAATGGCTGAGGCTGACGATAT 1401  
Db 11357 TAAAAAATGGCTGAGGCTGACGATAT 11382

RESULT 6

ABK89343  
ID ABK89343 standard; DNA; 3450 BP.  
XX  
AC ABK89343;  
XX  
XX  
DT 21-OCT-2002 (first entry)  
XX  
DE Plasmid pBCSACT2 DNA.  
XX  
XX Arsenate reductase; ArsC; cyclic; circular; ds; antimony; zinc;  
KW cadmium; phytochelatin biosynthetic enzyme; arsenate; arsenite; cobalt;  
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;  
KW industrial waste; phytoremediation.  
XX  
OS Synthetic.  
XX  
PN WO200248335-A2.  
XX  
PD 20-JUN-2002.  
XX  
PF 13-DEC-2001; 2001WO-US048105.  
XX  
PR 13-DEC-2000; 2000US-0255001P.  
PR 22-JUN-2001; 2001US-0300525P.  
XX

PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX Meagher RB, Li Y;  
XX WPI; 2002-583507/62.  
XX  
XX Recombinant DNA molecules for producing transgenic plants tolerant to  
PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of  
PT contaminated soil or water, encodes arsenate reductase coding sequence.  
XX  
PS Disclosure; Page 75-77; 131pp; English.  
XX  
CC The invention relates to a nucleic acid molecule comprising a portion  
CC encoding an arsenate reductase coding sequence and a plant-expressible  
CC transcription regulatory sequence, the coding sequence being operably  
CC linked to the transcription regulatory sequence. The nucleic acid is  
CC useful for producing a plant which is resistant to at least one metal ion  
CC such as a divalent cadmium ion or antimony, by introducing the nucleic  
CC acid molecule into a plant cell or into plant tissue, selecting for the  
CC presence of the nucleic acid molecule to produce a transgenic plant cell  
CC or plant tissue and regenerating a plant from the transgenic plant cell  
CC or plant tissue. The method further comprises introducing at least one  
CC plant-expressible phytochelatin biosynthetic enzyme coding sequence into  
CC the cell to produce an arsenate metal ion resistant plant. The arsenate  
CC reductase coding sequence is expressed under the control of a plant  
CC promoter which directs expression in the above ground plant part. The  
CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,  
CC antimony, arsenate and arsenite ions. The transgenic plant is useful for  
CC bioremediation of arsenate and/or cadmium-contaminated environments,  
CC including soil, sediments, mine tailings, water, industrial waste,  
CC groundwater and air. The transgenics are also useful for revegetation of  
CC soils contaminated with metal ions and for removing and sequestering  
CC these ions from water, wastewater and aqueous environments. The plants  
CC are also useful for phytoremediation of contaminated soil, sediment,  
CC water and mine tailings. This sequence represents plasmid pBCSACT2 DNA,  
CC used in the scope of the invention  
XX

SQ Sequence 3450 BP; 946 A; 694 C; 764 G; 1046 T; 0 U; 0 Other;

Query Match 89.9%; Score 1298.4; DB 6; Length 3450;  
Best Local Similarity 98.1%; Pred. No. 1.4e-303;  
Matches 1376; Conservative 0; Mismatches 18; Indels 8; Gaps 6;

QY 1 ATTATGATCTCAATACATGATGATACATATCTCATCTAGATCTAGGTATATGTAAG 60  
Db 9 ATGCTGATCTCAATACATGATGATACATATCTCATCTAGATCTAGGTATATGTAAG 68  
QY 61 AAAGTTTGGCAATATGNNACGACAAATGGCTACACTCGATGTAATGGTATCTCAAC 120  
Db 69 AAAGTTTGGCAATATGGCAGCAGCAAAATGGCTAGACTCGATGTAATGGTATCTCAAC 128  
QY 121 TCAACATTATATCTTACCAACATTAGTTAG-CAAAATTTAAACAACA-TTTTATGT 178  
Db 129 TCAACATTATATCTTACCAACATTAGTTAGACAAAATTTAAACAACA-TTTTATGT 188  
QY 179 ATGCAAGAGTCAGCATATGTATAATTTGATTGAGATCGTTTGGACGAGTTCGGATGTAGT 238  
Db 189 ATGCAAGAGTCAGCATATGTATAATTTGATTGAGATCGTTTGGACGAGTTCGGATGTAGT 248  
QY 239 AGTAGCCATTATTTAATGATGATACATCTAATCTGTGAATGTG-ATATGATGAACAATTGAT 297  
Db 249 AGTAGCCATTATTTAATGATGATACATCTAATCTGTGAATGTGATGAATGAACAATTGAT 308  
QY 298 CTTATGTTATAAATATCCATAAACAACATCATGAAGACACTTCTTTCCAGGCTCTGAATT 357  
Db 309 CTTATGTTATAAATATCCATAAACAACATCATGAAGACACTTCTTTCCAGGCTCTGAATT 368  
QY 358 AATTATGATACAATTTCTAATAGAAAAAGCAATTTAAATTTACGTTGAATTTGATGAATCTAA 417  
Db 369 AATTATGATACAATTTCTAATAGAAAAAGCAATTTAATTTACGTTGAATTTGATGAATCTAA 428  
QY 418 TTGAACAGGCCAACCAACGACGAGACTAAACGTTCCCTGGATTTGACTCGGTTTAAGTTAAC 477

Db 429 TTGAACAAGCCAAACACGACGACGACTAACGTTGCTGGAATGACTCGGTTAAGTTAAC 488  
Qy 478 CACTAAAAAAGCGAGCTGCTATGTAACACGCGGATCGAGCAGGTCACAGTCATGAAGCC 537  
Db 489 CACTAAAAAAGCGAGCTGCTATGTAACACGCGGATCGAGCAGGTCACAGTCATGAAGCC 548  
Qy 538 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAAGTTAAAAATTAAGTTAA 597  
Db 549 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAAGTTAAAAATTAAGTTAA 608  
Qy 598 CACGAGGAAAAA-GCTGTCTGACAGCAGGTCAAGTATCTTTTACCTGTGTCGAAAAAGA 656  
Db 609 CACGAGGAAAAAGGCTGTCTGACAGCAGGTCAAGTATCTTTTACCTGTGTCGAAAAAGA 668  
Qy 657 TTTCTGTCTGTCTGATTTTATTTATTTTAAAGGCGGAAATAAAGTTGTAAGAGATA 716  
Db 669 TTTCTGTCTGTCTGATTTTATTTATTTTAAAGGCGGAAATAAAGTTGTAAGAGATA 728  
Qy 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGCTCTCGTTGCTCTCC 776  
Db 729 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGCTCTCGTTGCTCTCC 788  
Qy 777 TCACCTTTCATAGCCGCTTTTGAATCTCCGCGACCTTGACAGAGAAACAAGAAAGAGA 836  
Db 789 TCACCTTTCATAGCCGCTTTTGAATCTCCGCGACCTTGACAGAGAAACAAGAAAGAGA 848  
Qy 837 CTAAGAGAGAAAGTAAGAGATTAATCCAGGAGATTCATTTCTCGTTTGAATCTTCTCAA 896  
Db 849 CTAAGAGAGAAAGTAAGAGATTAATCCAGGAGATTCATTTCTCGTTTGAATCTTCTCAA 908  
Qy 897 TCTCATCTTCTTCTCCGCTTTTCTTCCAAAGGTAATAGGAACCTTTCGGAATCTACTTT 956  
Db 909 TCTCA---TCTTCTCCGCTTTTCTTCCAAAGGTAATAGGAACCTTTCGGAATCTACTTT 965  
Qy 957 ATTGCTGATCTCGATCTGTTTCTCAATTTCTTGATGATCTGGAATCTGTTAAATTT 1016  
Db 966 ATTGCTGATCTCGATCTGTTTCTCAATTTCTTGATGATCTGGAATCTGTTAAATTT 1025  
Qy 1017 GATCTGTGAACCTCCACATAATCTTTGTTTCTAGAACTGATCTGAATCTGAATCTGACCGAT 1076  
Db 1026 GATCTGTGAACCTCCACATAATCTTTGTTTCTAGAACTGATCTGAATCTGAATCTGACCGAT 1085  
Qy 1077 CAGTTAGCTCGATTTAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCATGTTTC 1136  
Db 1086 CAGTTAGCTCGATTTAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCATGTTTC 1145  
Qy 1137 ATGTTACCTGGGAAATGATTTGATATGTAATGAAATCTGAAATCTGTTGAAAGTTAGATT 1196  
Db 1146 ATGTTACCTGGGAAATGATTTGATATGTAATGAAATCTGAAATCTGTTGAAAGTTAGATT 1205  
Qy 1197 GAATCTGAACCTGCAATGTTAGATTGAATCTGACACCTGTTTAA-GTTAGATGAAGTT 1255  
Db 1206 GAATCTGAACCTGCAATGTTAGATTGAATCTGACACCTGTTTAAAGGTTAGATGAAGTT 1265  
Qy 1256 TGTGTATAGATTCTTGGAAACCTTAGGATTTGTAGTGTCTAGCTTGAACAGAAAGCTAT 1315  
Db 1266 TGTGTATAGATTCTTGGAAACCTTAGGATTTGTAGTGTCTAGCTTGAACAGAAAGCTAT 1325  
Qy 1316 TTTCTGATCAATCAGGGTTTATTTGATCTGATGAAATCTTTTGTGTTTGTGAGCTCA 1375  
Db 1326 TTTCTGATCAATCAGGGTTTATTTGATCTGATGAAATCTTTTGTGTTTGTGAGCTCA 1385  
Qy 1376 TAAAAATGGCTGAGGCTGACG 1397  
Db 1386 TAAACCATGGCAATCCCGGACG 1407

RESULT 7

ADN61599

ID ADN61599 standard; DNA; 7474 BP.

XX

AC ADN61599;

XX

DT 29-JUL-2004 (first entry)  
XX COR102 event nucleotide sequence SEQ ID NO:7.  
XX COR102 motif event; insect resistant plant; VIP3A; detection; plant;  
KW COR102 event; insecticidal; Heliothis; Helicoverpa; Spodoptera;  
KW cotton boll worm; gene; ds; insect resistant transgenic cotton event.  
OS Synthetic.  
XX WO2004039986-A1.  
XX 13-MAY-2004.  
XX 23-OCT-2003; 2003WO-EP011725.  
XX 29-OCT-2002; 2002GB-00025129.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
PA Ellis DM, Negroto DV, Shi L, Shotkoski FA, Thomas CR;  
PI WPI; 2004-390327/36.  
DR Novel polynucleotide comprising contiguous nucleotides of motif  
PT designated COR102 event, useful in detecting insect resistant plant  
PT material derived from the event.  
XX Claim 16; SEQ ID NO 7; 64pp; English.  
XX The present invention describes a polynucleotide (I) comprising at least  
CC 17 contiguous nucleotide of the insect resistant transgenic cotton  
CC (COR102) motif event having the 26 nucleotide sequence of SEQ ID NO:1 or  
CC SEQ ID NO:2. Also described: (1) an insect resistant plant (II)  
CC comprising a VIP3A protein and (1); (2) detecting (M1) plant material  
CC derived from the COR102 event; and (3) a kit of parts comprising a unit  
CC for (M1). (I) and (M1) are useful for detecting a plant material derived  
CC from the COR102 event. (I) enables efficient detection of a plant  
CC material derived from the COR102 event. (II) has an insecticidal effect  
CC on insects from one or more species chosen from Heliothis sp.,  
CC Helicoverpa sp. and Spodoptera sp.. (II) has enhanced self-defence  
CC mechanism against infestation by pest insects such as Helicoverpa zea  
CC (cotton boll worm), and so reduces the number of insecticide sprays  
CC during the cultivation of (II) compared to non-transgenic cotton plant of  
CC the same variety and yield loss through insect pests in kept at a minimal  
CC level. The present sequence represents a COR102 event nucleotide  
CC sequence, which is used in the exemplification of the present invention.  
XX SQ Sequence 7474 BP; 2223 A; 1777 C; 1647 G; 1827 T; 0 U; 0 Other;  
Query Match 89.9%; Score 1298.4; DB 12; Length 7474;  
Best Local Similarity 98.8%; Pred. No. 1.7e-303;  
Matches 1370; Conservative 0; Mismatches 8; Indels 8; Gaps 6;  
Qy 1 ATTATGATCTCAATATACATATCTCATCTAGATCTAGCTATCATTTATGTAAG 60  
Db 3228 ATTATGATCTCAATATACATATCTCATCTAGATCTAGCTATCATTTATGTAAG 3287  
Qy 61 AAAGTTTTGACGAATATGNNACGACAAAATGGCTACCTCGATGTAATGGTATCTCAAC 120  
Db 3288 AAAGTTTTGACGAATATGGCAGCAAAAATGGCTAGACTCGATGTAATGGTATCTCAAC 3347  
Qy 121 TCAACATTATCTTATATACCAACATTAGTAG-CAAAATTTAAACAACTA-TTTTATGT 178  
Db 3348 TCAACATTATCTTATATACCAACATTAGTAGCAAAAATTTAAACAACTATTTTATGT 3407  
Qy 179 ATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTTCAGAGTTCGATGTAGT 238  
Db 3408 ATGCAAGAGTCAGCATATGTATAATTTGATTCAGAAATCGTTTTCAGAGTTCGATGTAGT 3467  
Qy 239 AGTAGCCATTATTTAATGTACATCTAATCGTGAATAGTG-ATATGATGAACATTTAT 297  
Db 3468 AGTAGCCATTATTTAATGTACATCTAATCGTGAATAGTGATGATGAACATTTAT 3527



QY 121 TCACATTATACCTTATACCAACATTTAGTAGT-CAAAAATTTAAACAACCTA-TTTTATGT 178  
|||||  
Db 4734 TCACATTATACCTTATACCAACATTTAGTAGAGCAAAAATTTAAACAACCTA-TTTTATGT 4793  
QY 179 ATGCAAGAGTCAGCATATGATTAATTTAGATTCAGAAATCGTTTTCAGAGTTCGGATGTAGT 238  
Db 4794 ATGCAAGAGTCAGCATATGATTAATTTAGATTCAGAAATCGTTTTCAGAGTTCGGATGTAGT 4853  
QY 239 AGTAGCCATTATTTAATGTPACATCTAATTCGTTGAATAGTG-ATATGATGAACATTTGTAT 297  
Db 4854 AGTAGCCATTATTTAATGTPACATCTAATTCGTTGAATAGTG-ATATGATGAACATTTGTAT 4913  
QY 298 CTTATTGTATAAATATCCATAAACACATCATCAAAAGACACTTTTCTTTCAGGCTCTCAATT 357  
Db 4914 CTTATTGTATAAATATCCATAAACACATCATCAAAAGACACTTTTCTTTCAGGCTCTCAATT 4973  
QY 358 AATTATGATPACAAATCTTAATAGAAAACGAATTAATTTAGTTGAATTTGATGAATCTAA 417  
Db 4974 AATTATGATPACAAATCTTAATAGAAAACGAATTAATTTAGTTGAATTTGATGAATCTAA 5033  
QY 418 TTGAACAGCCAAACACGACGAGGACTAACGTTGCTCGATTGACTCGGTTTAAAGTTAAC 477  
Db 5034 TTGAACAGCCAAACACGACGAGGACTAACGTTGCTCGATTGACTCGGTTTAAAGTTAAC 5093  
QY 478 CACTAAAAAACGGAGCTCTCATGTAAACACGCGGATCGAGCAGGTCAACAGTCATCAAGCC 537  
Db 5094 CACTAAAAAACGGAGCTCTCATGTAAACACGCGGATCGAGCAGGTCAACAGTCATCAAGCC 5153  
QY 538 ATCAAAGCAAAAGAACTAATCAAGGGGTGAGATTAATTTAGTTTAAAAATTTAGTTAA 597  
Db 5154 ATCAAAGCAAAAGAACTAATCAAGGGGTGAGATTAATTTAGTTTAAAAATTTAGTTAA 5213  
QY 598 CACGAGGGGAAA-GCTGTCTGACACGCCAGGTCACTGTTATCTTTTACCTGTGGTTCGAAATGA 656  
Db 5214 CACGAGGGGAAAAGGCTGTCTGACAGCCAGGTCACTGTTATCTTTTACCTGTGGTTCGAAATGA 5273  
QY 657 TTCTGTCTGTCTCGATTTTAAATTAATTTTGTAAAGGCCGAAAATAAAGTTGTAAGAGATA 716  
Db 5274 TTCTGTCTGTCTCGATTTTAAATTAATTTTGTAAAGGCCGAAAATAAAGTTGTAAGAGATA 5333  
QY 717 AACCCGCCATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCGTGTGCTCTCC 776  
Db 5334 AACCCGCCATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCGTGTGCTCTCC 5393  
QY 777 TCACATTGATCAGCGTTTTGATCTCCGCGACATTGACAGAGAAACAAGGAAGAGA 836  
Db 5394 TCACATTGATCAGCGTTTTGATCTCCGCGACATTGACAGAGAAACAAGGAAGAGA 5453  
QY 837 CTAAGAGAGAAAAGTAAGAGATAATCCAGAGAGATTCAATCTCCGTTTTTGAATCTTCTCAA 896  
Db 5454 CTAAGAGAGAAAAGTAAGAGATAATCCAGAGAGATTCAATCTCCGTTTTTGAATCTTCTCAA 5513  
QY 897 TCTCATCTTCTTCTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTTCTGGATCTACTTT 956  
Db 5514 TCTCA--TCTTCTTCCGCTCTTTCTTCCAAAGGTAATAGGAACCTTTCTGGATCTACTTT 5570  
QY 957 ATTTGCTGATCTCGATCTGTTTTCTCAATTTCTTGGATCTGGAATTCGTTAAATTT 1016  
Db 5571 ATTTGCTGATCTCGATCTGTTTTCTCAATTTCTTGGATCTGGAATTCGTTAAATTT 5630  
QY 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAACTGAATTTGACCGAT 1076  
Db 5631 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAACTGAATTTGACCGAT 5690  
QY 1077 CAGTTAGCTCGATTATAGTACAGAAATTTGCTTGACCTTGATGAGAGATCCATGTTTC 1136  
Db 5691 CAGTTAGCTCGATTATAGTACAGAAATTTGCTTGACCTTGATGAGAGATCCATGTTTC 5750  
QY 1137 ATGTTACCTCGGAAATGATTTCTATATGTAATTCGAATCTGAATCTCACTGTTGAAGTTAGATT 1196  
Db 5751 ATGTTACCTCGGAAATGATTTCTATATGTAATTCGAATCTGAATCTCACTGTTGAAGTTAGATT 5810  
QY 1197 GAATCTGAACACTGTCTCAATGTTAGATTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255

Db 5811 GAACTCGAACACTGTCAATGTTAGATTGAACTCGAACACTGTTTAAAGGTTAGATGAAGTT 5870  
QY 1256 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTAGTGTGTAACAGAAAGCTAT 1315  
Db 5871 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTAGTGTGTAACAGAAAGCTAT 5930  
QY 1316 TTTGATTCAATCAAGGGTTTATTTGACGTATTTGAATCTCTTTTGTGTGTTGCAAGCTCA 1375  
Db 5931 TTTGATTCAATCAAGGGTTTATTTGACGTATTTGAATCTCTTTTGTGTGTTGCAAGCTCA 5990  
QY 1376 TAAAAA 1381  
|||||  
Db 5991 TAAAAA 5996  
|||||  
RESULT 9  
ABK89342  
ID ABK89342 standard; DNA; 3408 BP.  
XX ABK89342;  
AC ABK89342;  
XX 21-OCT-2002 (first entry)  
DT Plasmid pATPCSACT2 DNA.  
DE Arsenate reductase; ArsC; cyclic; circular; ds; antimonate; zinc;  
KW cadmium; phytochelatin biosynthetic enzyme; arsenate; arsenite; cobalt;  
KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;  
KW industrial waste; phytoremediation.  
XX Synthetic.  
OS  
XX WO200248335-A2.  
PN 20-JUN-2002.  
XX  
PD 13-DEC-2001; 2001WO-US048105.  
XX  
PF 13-DEC-2000; 2000US-0255001P.  
PR 22-JUN-2001; 2001US-0300525P.  
XX  
PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX  
PI Meagher RB, Li Y;  
XX WPI; 2002-583507/62.  
DR  
XX Recombinant DNA molecules for producing transgenic plants tolerant to heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of contaminated soil or water, encodes arsenate reductase coding sequence.  
PS Disclosure; Page 69-71; 131pp; English.  
XX  
CC The invention relates to a nucleic acid molecule comprising a portion encoding an arsenate reductase coding sequence and a plant-expressible transcription regulatory sequence, the coding sequence being operably linked to the transcription regulatory sequence. The nucleic acid is useful for producing a plant which is resistant to at least one metal ion such as a divalent cadmium ion or antimonate, by introducing the nucleic acid molecule into a plant cell or into plant tissue, selecting for the presence of the nucleic acid molecule to produce a transgenic plant cell or plant tissue and regenerating a plant from the transgenic plant cell or plant tissue. The method further comprises introducing at least one plant expressible phytochelatin biosynthetic enzyme coding sequence into the cell to produce an arsenate metal ion resistant plant. The arsenate reductase coding sequence is expressed under the control of a plant promoter which directs expression in the above ground plant part. The plant is also resistant to cadmium, cobalt, copper, mercury, zinc, antimony, arsenate and arsenite ions. The transgenic plant is useful for bioremediation of arsenate and/or cadmium-contaminated environments, including soil, sediments, mine tailings, water, industrial waste, groundwater and air. The transgenics are also useful for revegetation of









DT	10-SEP-2001 (first entry)	Db	593	GAGTTCCGATGTAGTAGCCATTATTTAAATGTACATACATAATCGTGAATAGTATG	652
XX	Chimeric CamV-Act2 promoter including first intron of Act-2 gene.				
DE	DNA construct; structural gene; SG; non-translated region; NTR; actin;	Qy	284	ATGAACATCTTATCTTATTTATTAATATCATATAACACACATCATGAAGACACTTTCTT	343
XX	Act; elongation factor 1 alpha; EFlalpha; weed control; EPSP synthase;				
KW	glyphosate tolerance; glyphosate oxidoreductase; herbicide tolerance;	Db	653	ATGAACATCTTATCTTATTTATTAATATCATATAACACATCATGAAGACACTTTCTT	712
KW	PCR primer; Act2 promoter; Cauliflower mosaic virus; CamV; chimeric; ds.	Qy	344	TCAGGCTCTGAATTAATATGATACAATTTCTAATAGAAAAAGAAATTAATTAATTAAT	403
XX		Db	713	TCACGCTCTGAATTAATATGATACAATTTCTAATAGAAAAAGAAATTAATTAATTAAT	772
OS	Cauliflower mosaic virus.	Qy	404	TGTATGAATCTAATTTGAAGCAACCAACGACGAGACTAAACGTTGCGTGGATGACT	463
OS	Arabidopsis thaliana.	Db	773	TGTATGAATCTAATTTGAAGCAACCAACGACGAGACTAAACGTTGCGTGGATGACT	832
XX	Chimeric.				
FH	Key				
FT	promoter	Qy	464	CGTTTAAAGTTAAACCACTAAATAAAGCGAGCTCATCTAAACGCGATCGAGCAGGTC	523
FT	Location/Qualifiers	Db	833	CGTTTAAAGTTAAACCACTAAATAAAGCGAGCTCATCTAAACGCGATCGAGCAGGTC	892
FT	1..523				
FT	/*tag= a	Qy	524	ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATAGTT	583
FT	/note= "CamV promoter"	Db	893	ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATAGTT	952
FT	534..1742				
FT	/*tag= b	Qy	584	TAAAAATTAAGTTAAACCAAGGGGAAAAGGCTGCTGACAGCCAGGTCACGTTATCTTACC	642
FT	/note= "Arabidopsis thaliana Act2 promoter, intron and 5'	Db	953	TAAAAATTAAGTTAAACCAAGGGGAAAAGGCTGCTGACAGCCAGGTCACGTTATCTTACC	1012
FT	UTR region"	Qy	643	TGTGTCGAAATGATTCGTGCTCGATTTTAAATTTTTCGAAAGGCCGAAATAA	702
XX		Db	1013	TGTGTCGAAATGATTCGTGCTCGATTTTAAATTTTTCGAAAGGCCGAAATAA	1072
PN	WO200144457-A2.	Qy	703	AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG	762
XX	21-JUN-2001.	Db	1073	AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG	1132
XX	12-DEC-2000; 2000WO-US033633.	Qy	763	TCTCGTTGCTCCTCACTTTTCATCAGCGGTTTTGAAATCTCGGCGACTTGACAGAGA	822
XX	16-DEC-1999; 99US-0171173P.	Db	1133	TCTCGTTGCTCCTCACTTTTCATCAGCGGTTTTGAAATCTCGGCGACTTGACAGAGA	1192
XX	(MONS ) MONSANTO CO.	Qy	823	AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT	882
PA	Fincher KL, Flasiniski S, Wilkinson JQ;	Db	1193	AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT	1252
PI	WPI; 2001-408480/43.	Qy	883	TGAATCTTCTCAATCTCATCTTCTTCTCCGCTCTTTCTTTTCCAAGGTAAATAGAACTT	942
DR	Novel DNA construct for controlling weeds, containing Arabidopsis actin	Db	1253	TGAATCTTCTCAATCTCA---TCTTCTTCCGCTCTTCTTTTCCAAGGTAAATAGAACTT	1309
XX	promoter sequences and elongation factor 1-alpha promoter sequences	Qy	943	TCTGATCTACTTTATTTATTTGCTGGATCTCGATCTGTTTTTCTCAATTTCTTGAGATCTGG	1002
PT	operably linked to structural genes that function in crop plants.	Db	1310	TCTGATCTACTTTATTTATTTGCTGGATCTCGATCTGTTTTTCTCAATTTCTTGAGATCTGG	1369
XX		Qy	1003	AATTGCTTTTAAATTTGGATCTGTGAACCTCCCAATAATCTTTTGGTTTTTACTAGAATCGAT	1062
PS	Claim 14; Page 100-101; 101pp; English.	Db	1370	AATTGCTTTTAAATTTGGATCTGTGAACCTCCCAATAATCTTTTGGTTTTTACTAGAATCGAT	1429
XX	The present invention relates to DNA construct comprising an expression	Qy	1063	CTAAGTTACCCGATCAGTTAGCTCGATTTATAGTACCAGAAATTTGGCTTTGACTTGATGG	1122
CC	cassette containing a promoter DNA sequence operably linked to a	Db	1430	CTAAGTTACCCGATCAGTTAGCTCGATTTATAGTACCAGAAATTTGGCTTTGACTTGATGG	1489
CC	structural gene (SG) encoding an agronomically useful protein, and a 3'	Qy	1123	AGAGATCCATGTTTCAATTTAGTTACTCGGAAATGATTTGTATATGTGAATTAATTAATTAAT	1182
CC	non-translated region (NTR) that functions in plants to cause the	Db	1490	AGAGATCCATGTTTCAATTTAGTTACTCGGAAATGATTTGTATATGTGAATTAATTAATTAAT	1549
CC	addition of polyadenylated nucleotides to the 3' end of the RNA sequence,	Qy	1183	GTGGAAGTTAGATTAATCTGAACACTCTCAATGTTAGATTAATCTGAACACTCTTTAA	1242
CC	operably linked to SG. The present invention further relates to plant	Db	1550	GTGGAAGTTAGATTAATCTGAACACTCTCAATGTTAGATTAATCTGAACACTCTTTAA	1609
CC	expression constructs that comprise Arabidopsis actin (Act) promoter	Qy	1243	GTGGAAGTTAGATTAATCTGAACACTCTCAATGTTAGATTAATCTGAACACTCTTTAA	1302
CC	sequences Act1a, Act1b, Act2, Act3, Act7, Act8, Act11, Act12, elongation	Db	1610	GTGGAAGTTAGATTAATCTGAACACTCTCAATGTTAGATTAATCTGAACACTCTTTAA	1669
CC	factor 1 alpha (EFlalpha) promoter sequence, fragments and cis elements	Qy	1303	AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGACGTGATTTGAACCTTTTGTG	1362
CC	derived from these promoters operably linked to heterologous structural	Db	1670	AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGACGTGATTTGAACCTTTTGTG	1729
CC	gene sequences that function in crop plant cells. The DNA constructs are				
CC	useful for controlling weeds, by providing a crop plant transformed with				
CC	a DNA construct comprising promoter DNA sequence that is functional in a				
CC	plant cell, a glyphosate tolerance gene (which is an EPSP synthase gene				
CC	or a glyphosate oxidoreductase gene) or a herbicide tolerance gene. The				
CC	present sequence is chimeric Cauliflower mosaic virus (CamV)-Act2				
CC	promoter including first intron of Arabidopsis thaliana Act2 gene.				
CC	(Updated on 11-SEP-2003 to standardise OS field)				
XX	Sequence 1742 BP; 527 A; 327 C; 362 G; 526 T; 0 U; 0 Other;				
SQ					
	Query Match				
	Best Local Similarity 99.3%; Score 1174; DB 4; Length 1742;				
	Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;				
Qy	164 CAACATTTTATGATGACAGAGTCAGCATATGATATTAATTCATTCAGAAATCGTTTTCAC				223
Db	533 CAACATTTTATGATGACAGAGTCAGCATATGATATTAATTCATTCAGAAATCGTTTTCAC				592
Qy	224 GAGTTCCGATGTAGTAGCCATTATTTAAATGTACATACATAATCGTGAATAGTATG				283





Db 944 AGAGATCCATGTTTCATGTTACCTGGGAATGATTTGTATATGGAATGAAATCGAACT 1003  
QY 1183 GTTGAAGTTAGATGTAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1242  
Db 1004 GTTGAAGTTAGATGTAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1063  
QY 1243 GTTAGATGAAGTTTGTGTAGATGTTCTCGAACCTTAGGATTTGAGTGTGCTAGTTG 1302  
Db 1064 GTTAGATGAAGTTTGTGTAGATGTTCTCGAACCTTTAGGATTTGAGTGTGCTAGTTG 1123  
QY 1303 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTG 1362  
Db 1124 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTG 1183  
QY 1363 TGTTCGAGCTCATAAAAAATGCTGAGGCTGACGATATTCACCAATCGTGTGTGACAA 1422  
Db 1184 TGTTCGAGCTCATAAAAAATGCTGAGGCTGATGATATTCACCAATCGTGTGTGACAA 1243  
QY 1423 TGGTACTGGAATGGT 1437  
Db 1244 TGGTACTGGAATGGT 1258

RESULT 15  
AADO1288  
ID AADO1288 standard; DNA; 15676 BP.  
XX AC AADO1288;  
XX DT 12-OCT-2000 (first entry)  
XX DE Arabidopsis transformation binary vector, pAct2-bin.  
XX KW Matrix Attachment Region; MAR; scaffold attachment region.  
XX KW Arabidopsis transformation binary vector; pAct2-bin; gene expression;  
XX KW transgenic organism; ds.  
XX OS Synthetic.  
XX PN WO200032800-A1.  
XX PD 08-JUN-2000.  
XX PF 30-NOV-1999; 99WO-US028123.  
XX PR 01-DEC-1998; 98US-0110437P.  
XX PA (DOWC ) DOW AGROSCIENCES LLC.  
XX PI Van Der Geest AHM, Ainley WM, Cowen NM, Welter ME, Woosley AT;  
XX WPI; 2000-412345/35.  
XX An isolated DNA molecule for use as a matrix attachment region to  
PT increase expression of genes introduced in transformed plants comprises a  
PT 298 base pair sequence described in the specification.  
XX Example 3; Page 55-59; 73pp; English.

XX The patent discloses a DNA molecule, useful as matrix attachment region  
CC (MAR) or scaffold attachment region, to increase the expression of genes  
CC introduced in transformed plants. MARs are located in non-transcribed  
CC regions of genes and form the physical boundaries of individual DNA  
CC loops. They are rich in adenosine and thymine bases and contain certain  
CC conserved sequence elements and structural features. They are about 300-  
CC 2000 bp in length. Increased levels of expression of DNA introduced into  
CC plants can be achieved by use of MAR. They can also reduce the position  
CC effect in transgenic organisms. The present DNA sequence is the  
CC Arabidopsis transformation binary vector pAct2-bin, that contains a  
CC Act2/GUS/nos cassette (Act2 transcription initiation region/GUS  
CC structural gene/nos 3' untranslated region), 19S/NPTII/orf25polyA as a  
CC selectable marker and 35S/GFP/nos as an independent reporter gene. This

CC binary vector is used to construct other vectors, used to test the two  
CC orientations of the artificial MAR dimers in Arabidopsis  
XX  
SQ Sequence 15676 BP; 4226 A; 3741 C; 3838 G; 3868 T; 0 U; 3 Other;  
Query Match 74.7%; Score 1078.8; DB 3; Length 15676;  
Best Local Similarity 98.3%; Pred. No. 2.7e-250;  
Matches 1154; Conservative 0; Mismatches 12; Indels 8; Gaps 6;  
QY 205 GATTCAGAAATCTTTTACGAGTTCGGATGTAGTAGTCCCATTTATTAATGTCATACAT 264  
Db 84 GAAATTCGGCTTGTGTTGACGAGTTCGGATGTAGTAGTCCCATTTATTAATGTCATACAT 143  
QY 265 AATCGTGAATAGTG-ATATGATGAAACATTTGATCTTATTTGATATAAATATCCATAAACAC 323  
Db 144 AATCGTGAATAGTGATGATGAAACATTTGATCTTATTTGATATAAATATCCATAAACAC 203  
QY 324 ATCATGAAGACACATTTCTTCAGGGTCTGAATTAATTAATGATACAAATCTTAATAGAAAA 383  
Db 204 ATCATGAAGACATTTCTTCACGGTCTGAATTAATTAATGATACAAATCTTAATAGAAAA 263  
QY 384 CGAATTAATACGTTGAAATTTGATGAATCTAAATTTGAACAGCCACACCCAGCAGAC 443  
Db 264 CGAATTAATACGTTGAAATTTGATGAATCTAAATTTGAACAGCCACACCCAGCAGAC 323  
QY 444 TAACGTTGCTGGATTGACTCGGTTTAAAGTTAAACCACTAAATAAAGCGAGCTGTCTATGTA 503  
Db 324 TAACGTTGCTGGATTGACTCGGTTTAAAGTTAAACCACTAAATAAAGCGAGCTGTCTATGTA 383  
QY 504 ACACGGGATCGAGCAGGTCAAGTCATGATGAGCCATCAAGCAAGAAAGAACTAATCCAAAG 563  
Db 384 ACACGGGATCGAGCAGGTCAAGTCATGATGAGCCATCAAGCAAGAAAGAACTAATCCAAAG 443  
QY 564 GG-TGAGATGATTAATTAAGTTTAAATAATAGTTTAAACAGGAGGAAAA-GCTCTCTGACAG 621  
Db 444 GGCTGAGATGATTAATTAAGTTTAAATAATAGTTTAAACAGGAGGAAAAAGCTCCCTGACAG 503  
QY 622 CCAGTCACTGTTATCTTTACCTGCTGGTGTGAAATGATTCGTTGCTGCTCGATTTTAATTAAT 681  
Db 504 CCAGTCACTGTTATCTTTACCTGCTGGTGTGAAATGATTCGTTGCTGCTCGATTTTAATTAAT 563  
QY 682 TTTTGAAGGCGGAAATTAAGTTTGAAGATTAACCCGCTATATAAATTCATATAT 741  
Db 564 TTTTGAAGGCGGAAATTAAGTTTGAAGATTAACCCGCTATATAAATTCATATAT 623  
QY 742 TTTCTCTCCGCTTTGAAATTTGCTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 801  
Db 624 TTTCTCTCTCCGCTTTGAAATTTGCTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 683  
QY 802 TCCGCGCACTTTCAGAGAGAGAAACAAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861  
Db 684 TCCGCGCACTTTCAG 743  
QY 862 CAGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCATCTCTCTCTCTCTCTCTCTCTCTCT 921  
Db 744 CAGAGATTCATTTCTCCGTTTGAATCTTCTCTCAATCTCTCACTCTCTCTCTCTCTCTCTCT 800  
QY 922 TTTCCAGGTAATAGGAACTTTCTGGATCTACTTTTATTTGCTGGATCTCGATCTCTGTTT 981  
Db 801 TTTCCAGGTAATAGGAACTTTCTGGATCTACTTTTATTTGCTGGATCTCGATCTCTGTTT 860  
QY 982 CTCAATTTCTTGAGATCTGGAAATTCGTTTAAATTT-GGATCTGTGAACCTCCACCTAAATC 1040  
Db 861 CTCAATTTCTTGAGATCTGGAAATTCGTTTAAATTTGGGATCTGTGAACCTCCACCTAAATC 920  
QY 1041 TTTTGGTTTACTAGATCGATCTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACCA 1100  
Db 921 TTTTGGTTTACTAGATCGATCTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACCA 980  
QY 1101 GAAATTCGCTTGACCTTGATCGAGAGATCCCATGTTTCAATGTTTACCTGGGAAATGATTTGTA 1160  
Db 981 GAAATTCGCTTGACCTTGATCGAGAGATCCCATGTTTCAATGTTTACCTGGGAAATGATTTGTA 1040

Qy	1161	TATGTGAATGAAATCTGAACTGTTCAGCTTAGATTGAATCTGAACTGTCAATGTTAG	1220
Db	1041	TATGTGAATGAAATCTGAACTGTTCAGCTTAGATTGAATCTGAACTGTCAATGTTAG	1100
Qy	1221	ATTGAATCTGAACTGTTTAA-GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTT	1279
Db	1101	ATTGAATCTGAACTGTTTAAAGTTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTT	1160
Qy	1280	AGGATTTGTAGTGTCTGCTGTTGAACAGAAAGCTATTTCTGATTCGAATCAGGGTTTATTT	1339
Db	1161	AGGATTTGTAGTGTCTGCTGTTGAACAGAAAGCTATTTCTGATTCGAATCAGGGTTTATTT	1220
Qy	1340	GACTGTATTGAACTCTTTTGTGTGTTTGCAGCT	1373
Db	1221	GACTGTATTGAACTCTTTTGTGTGTTTGCAGCT	1254

Search completed: May 27, 2005, 20:09:22  
 Job time : 855 secs

1	1174.8	81.4	1219	3	US-09-737-6988-9	Sequence 9, Appli
2	1174.8	81.4	1219	4	US-09-737-628A-9	Sequence 9, Appli
3	1174	81.3	1742	3	US-09-737-6988-30	Sequence 30, Appli
4	1174	81.3	1742	4	US-09-737-628A-30	Sequence 30, Appli
5	75.4	5.2	1271	3	US-09-737-6988-10	Sequence 10, Appli
6	75.4	5.2	1271	4	US-09-737-628A-10	Sequence 10, Appli
7	73.8	5.1	1800	3	US-09-737-6988-29	Sequence 29, Appli
8	73.8	5.1	1800	4	US-09-737-626A-29	Sequence 29, Appli
9	52.8	3.7	1141	4	US-09-806-7088-22	Sequence 22, Appli
10	49	3.4	19438	1	US-09-949-016-13699	Sequence 12699, A
11	48.6	3.4	731	1	US-08-451-405A-2	Sequence 2, Appli
12	48.2	3.3	95255	4	US-09-949-016-17067	Sequence 17067, A
13	47.8	3.3	5203	3	US-09-257-770-1	Sequence 1, Appli
14	47.2	3.3	109038	4	US-09-949-016-12199	Sequence 12199, A
15	47	3.3	1141	4	US-09-806-7088-22	Sequence 22, Appli
16	46.6	3.2	18773	4	US-09-949-016-14164	Sequence 14164, A
17	46.6	3.2	640681	4	US-09-790-988-1	Sequence 1, Appli
18	46.2	3.2	2448	1	US-08-526-964-2	Sequence 2, Appli
19	46.2	3.2	2448	2	US-08-946-617-2	Sequence 2, Appli
20	46.2	3.2	2448	3	US-09-031-897-2	Sequence 2, Appli
21	46.2	3.2	30820	4	US-09-949-016-17145	Sequence 17145, A
22	46	3.2	134987	4	US-09-949-016-15348	Sequence 15348, A
23	46	3.2	134987	4	US-09-949-016-15349	Sequence 15349, A
24	46	3.2	134987	4	US-09-949-016-15350	Sequence 15350, A
25	46	3.2	134987	4	US-09-949-016-15507	Sequence 15507, A
26	46	3.2	134987	4	US-09-949-016-15508	Sequence 15508, A
27	46	3.2	134987	4	US-09-949-016-15509	Sequence 15509, A

```
Db 301 |||||TAAGTTAACTAACCACTAAAAACCGAGCTGTCATGTAAACACGCGATCAGCAGGTC 360
Qy 524 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATAGTT 583
Db 361 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATAGTT 420
Qy 584 TAAAAATTAAGTTAAACACGAGGAAAAA-GCTGTCACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTAAGTTAAACACGAGGAAAAAGGCTGTCACAGCCAGGTCACGTTATCTTTACC 480
Qy 643 TGTGTCGAAATGATTCGTGTCGTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 702
Db 481 TGTGTCGAAATGATTCGTGTCGTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 540
Qy 703 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTG 762
Db 541 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTG 600
Qy 763 TCTCGTTGTCCTCCTCACCTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGA 822
Db 601 TCTCGTTGTCCTCCTCACCTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGA 660
Qy 823 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAAATCCAGGAGATTCATTCTCCGTTT 882
Db 661 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAAATCCAGGAGATTCATTCTCCGTTT 720
Qy 883 TGAATCTCTCAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
Db 721 TGAATCTCTCAATCTCTCA---TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 777
Qy 943 TCTGATCTACTTTTATTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1002
Db 778 TCTGATCTACTTTTATTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 837
Qy 1003 AATTCTGTTTAAATTTGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT 1062
Db 838 AATTCTGTTTAAATTTGATCTGTAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT 897
Qy 1063 CTAAGTTGACCATCAGTTAGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1122
Db 898 CTAAGTTGACCATCAGTTAGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 957
Qy 1123 AGAGATCCATGTTTACCTGCGGAAATGATTTGATGATGATGATGATGATGATGATGATG 1182
Db 958 AGAGATCCATGTTTACCTGCGGAAATGATTTGATGATGATGATGATGATGATGATGATG 1017
Qy 1183 GTTGAAGTTAGATTTGAATCTGAACACTGTCATGTTAGATTGAATCTGAACACTGTTTAA 1242
Db 1018 GTTGAAGTTAGATTTGAATCTGAACACTGTCATGTTAGATTGAATCTGAACACTGTTTAA 1077
Qy 1243 GTTAGATGAAGTTTGTGATAGATTTCTCGAAACCTTAGGATTTGATGTCGTACGTTG 1302
Db 1078 GTTAGATGAAGTTTGTGATAGATTTCTCGAAACCTTTAGGATTTGATGTCGTACGTTG 1137
Qy 1303 AACAGAAAGCTATTTCTGATCTCAATCAGGCTTTTATTTGATGATGATGATGATGATG 1362
Db 1138 AACAGAAAGCTATTTCTGATCTCAATCAGGCTTTTATTTGATGATGATGATGATGATG 1197
Qy 1363 TGTGTCAGCTCATAAAAAATG 1384
Db 1198 TGTGTCAGCAGACTCACCATG 1219
```

RESULT 2

```
US-09-737-626A-9
; Sequence 9, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiński, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911e1 Plant Expression Constructs
```

```
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron
US-09-737-626A-9
```

```
Query Match 81.4%; Score 1174.8; DB 4; Length 1219;
Best Local Similarity 98.7%; Pred. No. 6.7e-309;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;
```

```
Qy 164 CAACCTATTTTATGTTATGTAAGAGTCCAGATATGTAATTAATTAATTAATTAATTAAT 223
Db 1 CAACCTATTTTATGTTATGTAAGAGTCCAGATATGTAATTAATTAATTAATTAATTAAT 60
Qy 224 GAGTTCCGATGTAGTAGCCATTATTTAATGTACATACATACTAATCGTGAATAGTCATATG 283
Db 61 GAGTTCCGATGTAGTAGCCATTATTTAATGTACATACATACTAATCGTGAATAGTCATATG 120
Qy 284 ATGAAACATGTATCTTATTTATTAATAATATCATATAACACATCATGATAAGACATTTCTT 343
Db 121 ATGAAACATGTATCTTATTTATTAATAATATCATATAACACATCATGATAAGACATTTCTT 180
Qy 344 TCAGGCTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 403
Db 181 TCAGGCTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Qy 404 TGTATGAATCTAAATTTGAACCAAGCCAAACAGCAGGAGCTAAACGTTGCTGGATTGACT 463
Db 241 TGTATGAATCTAAATTTGAACCAAGCCAAACAGCAGGAGCTAAACGTTGCTGGATTGACT 300
Qy 464 CGTTTAAAGTTAAACCACTAAACCAAGCGAGCTGTCATGTAAACACGCGATCAGCAGGTC 523
Db 301 CGTTTAAAGTTAAACCACTAAACCAAGCGAGCTGTCATGTAAACACGCGATCAGCAGGTC 360
Qy 524 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAATTAAT 583
Db 361 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAATTAAT 420
Qy 584 TAAAAATTAAGTTAAACACGAGGAAAAA-GCTGTCACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTAAGTTAAACACGAGGAAAAAGGCTGTCACAGCCAGGTCACGTTATCTTTACC 480
Qy 643 TGTGTCGAAATGATTCGTGTCGTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 702
Db 481 TGTGTCGAAATGATTCGTGTCGTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 540
Qy 703 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTG 762
Db 541 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTG 600
Qy 763 TCTCGTTGTCCTCCTCACCTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGA 822
Db 601 TCTCGTTGTCCTCCTCACCTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGA 660
Qy 823 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAAATCCAGGAGATTCATTCTCCGTTT 882
Db 661 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAAATCCAGGAGATTCATTCTCCGTTT 720
Qy 883 TGAATCTCTCAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
Db 721 TGAATCTCTCAATCTCTCA---TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 777
```



Qy	943	TCTCGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCCCTTGAGATCTGG	1002
Db	778	TCTCGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCCCTTGAGATCTGG	837
Qy	1003	AAATCGTTTAAATTTGGATCTGTGAACCTCCAACTAAATCTTTTGGTTTTTACTTAGAATCGAT	1062
Db	838	AAATCGTTTAAATTTGGATCTGTGAACCTCCAACTAAATCTTTTGGTTTTTACTAGNATCGAT	897
Qy	1063	CTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACCAAGAAATTTGGCTTGACCTTGATGG	1122
Db	898	CTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACCAAGAAATTTGGCTTGACCTTGATGG	957
Qy	1123	AGAGATCATCTTCACTGTTACTCGGAAATGATTTGTATGTCAATTTGAAATCTGAACCT	1182
Db	958	AGAGATCATCTTCACTGTTACTCGGAAATGATTTGTATGTCAATTTGAAATCTGAACCT	1017
Qy	1183	GTTCAAGTTAGATTTGAATCTGAACACTGTCAATCTTAGATTGAATCTGAACACTGTTTAA	1242
Db	1018	GTTCAAGTTAGATTTGAATCTGAACACTGTCAATCTTAGATTGAATCTGAACACTGTTTAA	1077
Qy	1243	GTTAGATGAAGTTTGTGTATAGATCTTCGAAACCTTAGAATTTGTAGTGTCTGACGTTG	1302
Db	1078	GTTAGATGAAGTTTGTGTATAGATCTTCGAAACCTTAGAATTTGTAGTGTCTGACGTTG	1137
Qy	1303	AACAGAAAGCTATTTCTTGATTTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1362
Db	1138	AACAGAAAGCTATTTCTTGATTTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1197
Qy	1363	TGTTTGACGCTCATAAAAAATG	1384
Db	1198	TGTTTGACGAGACTCACCATG	1219

### RESULT 3

```

US-09-737-698B-30
; Sequence 30, Application US/09737698B
; Patent No. 6462258
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6462258el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)C
; CURRENT APPLICATION NUMBER: US/09/737,698B
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/171,173
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 polynucleotides + Act2 ir
; OTHER INFORMATION: tro
; US-09-737-698B-30

```

	Query Match	81.3%	Score 1174;	DB 3;	Length 1742;
	Best Local Similarity	99.3%	Pred. No. 1.3e-308;		
	Matches 1201; Conservative	0;	Mismatches 5;	Indels 4;	Gaps 2
Qy	164 CAACCTATTTTATGTAATCGAAGAGTCAGCATATGTATAATTGAATTCAGAATCGTTTTGC				223
Dd	533 CAACCTATTTTATGTAATCGAAGAGTCAGCATATGTATAATTGAATTCAGAATCGTTTTGC				592
Qy	224 GAGTTCGGATGTAGTAGTGCCATTATTTAAATGTAACAATACTAATCGTGAATAGTGATATG				283
Dd	593 GAGTTCGGATGTAGTAGTAGCCATTATTTAAATGTAACAATACTAATCGTGAATAGTGATATG				652
Qy	284 ATGAACAATTCGTACTTTATTTGTAATAAATATCCATAAACACATCATGAAGAACACATTTCTT				343

Db	653	ATGAAACATTGTATCTTATTGTATTAATATCCATAAACAACATCATGATGAAGACACTTTCTTT	712
Qy	344	TCAGGGTCTGAAATTAATATATGATACAAATCTTAATAGAAAAAGAAATTAATTAAGCTTGAAT	403
Db	713	TCAGGTCTGAAATTAATATATGATACAAATCTTAATAGAAAAAGAAATTAATTAAGCTTGAAT	772
Qy	404	TGTATGAATCTTAATTTGAAACAAGCCAAACACGACGAGGACTAAACGTTTGCCTGGAATGACT	463
Db	773	TGTATGAATCTTAATTTGAAACAAGCCAAACACGACGAGGACTAAACGTTTGCCTGGAATGACT	832
Qy	464	CGGTTTAAGTTAAACCACTAAAAAAACGGAGCTGTCAATGTAACACGCGGATCGAGCAGGTC	523
Db	833	CGGTTTAAGTTAAACCACTAAAAAAACGGAGCTGTCAATGTAACACGCGGATCGAGCAGGTC	892
Qy	524	ACAGTCAATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATGAATTAATAGATT	583
Db	893	ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATGAATTAATAGATT	952
Qy	584	TAAAAATTAGTTAAACAGAGGGGAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTTACC	642
Db	953	TAAAAATTAGTTAAACAGAGGGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTTACC	1012
Qy	643	TGTGGTCGAAATAGATTGCTGTCTGTCCGATTTTAATTAATTTTTTTTGAAGGCCGAAATAA	702
Db	1013	TGTGGTCGAAATAGATTGCTGTCTGTCCGATTTTAATTAATTTTTTTTGAAGGCCGAAATAA	1072
Qy	703	AGTTGTAAGAGATAAAACCCGCCCTATATAAATTCATATAATTTTCTCTCCCGCTTTGAAATTG	762
Db	1073	AGTTGTAAGAGATAAAACCCGCCCTATATAAATTCATATAATTTTCTCTCCCGCTTTGAAATTG	1132
Qy	763	TCTCGTTGTCCTCTCACTTTTCATACAGCCGTTTGAATCTCCGGCGACTTGAACAGAGAAG	822
Db	1133	TCTCGTTGTCCTCTCACTTTTCATACAGCCGTTTGAATCTCCGGCGACTTGAACAGAGAAG	1192
Qy	823	AACAAGGAAGAACTAAGAGAGAAAGTAGAGATAATCCAGAGATTCATTCTCCGTTT	882
Db	1193	AACAAGGAAGAACTAAGAGAGAAAGTAGAGATAATCCAGAGATTCATTCTCCGTTT	1252
Qy	883	TGAATCTTCCTCAATCTCACTTTCTTCTCCGCTCTTTTCTTCCAAAGGTAAATAGGAACCTT	942
Db	1253	TGAATCTTCCTCAATCTCA---TCTTCTCCGCTCTTTCTTCCAAAGGTAAATAGGAACCTT	1309
Qy	943	TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTGTGAGATCTGG	1002
Db	1310	TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTTTCTCAATTTCTTGTGAGATCTGG	1369
Qy	1003	AATTCGTTTAAATTTGGATCTGTGAACCTCACTAAATCTTTTGGTTTACTAGAACTCGAT	1062
Db	1370	AATTCGTTTAAATTTGGATCTGTGAACCTCACTAAATCTTTTGGTTTACTAGAACTCGAT	1429
Qy	1063	CTAAGTTGACCGATCAGTTAGCTCGAATTAAGCTTACCCAGAAATTTGGCTTGACCTTGATGG	1122
Db	1430	CTAAGTTGACCGATCAGTTAGCTCGAATTAAGCTTACCCAGAAATTTGGCTTGACCTTGATGG	1489
Qy	1123	AGAGATCCATGTTTCAATGTTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACT	1182
Db	1490	AGAGATCCATGTTTCAATGTTTACCTGGGAAATGATTTGTATATGTAATGAAATCTGAACT	1549
Qy	1183	GTTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA	1242
Db	1550	GTTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA	1609
Qy	1243	GTTTGAAGTTAGTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGTCGTTG	1302
Db	1610	GTTTGAAGTTAGTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTCTGTCGTTG	1669
Qy	1303	AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTTATTTGACTGTATTTGAACTCTTTTGTG	1362
Db	1670	AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTTATTTGACTGTATTTGAACTCTTTTGTG	1729
Qy	1363	TGTTTGACG 1372	
Db	1730	TGTTTGACG 1739	



Matches	214;	Conservative	3;	Mismatches	194;	Indels	4;	Gaps	2
Qy	606	AAAAGCTGTTCTGCACAGCAGGTCA	CGTTA---TC	TTTTACCCTGGTGCGTAATGATTCGTG	662				
Db	471	AGAGGCTGTTCTGCACAGCCCATGTC	ACGTTATACTTTTTTCCG	TATGATCGAAATGATTCGTC	530				
Qy	663	TCGTGCAATTTTAATTATTTTTTGA	AAAGGCCGAAAAATAAGTTGT	TAAAGATATAAGATAAACCCG	722				
Db	531	TTTGYGAATTTAAATTATTTTCCA	AAATTTGAYGACTCTAAAGA	AAAAAATAAGTTTTTCA	590				
Qy	723	CCTATATAAATTTCATATATTTTCT	CCCCCGCTTTTGAAATTGCT	CGTTGCTCTCTCTCACTT	782				
Db	591	GATTAACCCGCTATATAAATAGTT	CAACACTCGGTTATTTCTTCT	CCCCCTCAAAGAT	650				
Qy	783	TCATCAGCCGTTTTTGAATCTCCG	CGACTTCACAGAGAAGAAACA	AGGAAGAAAGACTAAGA	842				
Db	651	TGCCTCGTGTCTTCAGCTTCATCG	CGCTTGCAATTTCCCGCGGATA	AAGAGAGAGAAAGA	710				
Qy	843	-GAGAAAGTAAGAGATTAATCCAG	GAGATTCAATCTCCGTTTGAAT	CTTCTCTCAATCTCA	902				
Db	711	GGAGAAAGAGTAGCCAGTTCTTC	CATCGTCGTGGTCTCTGTTTCT	TCTCTCGATCTCTCGA	770				
Qy	902	TCCTTCTTCTCCGCTCTTTCTTTC	CAAGGTAATAGGAACCTTCT	CGATCTACTTTATTG	961				
Db	771	TCCTCTGCTTTTGTCTTTTCCG	ATTAAAGTAATTAAACCTCCG	ATCTACTTGTCTTGTG	830				
Qy	962	CTGGATCTCGATCTTGTGTTTTCT	CAATTTCTTGAGATCTGGAAT	TCGTTTAAATTT	1016				
Db	831	TTGGATCTCGATTACGATTCTA	AGTTTACCTTCAAAGTGTGTT	CCCGATTGATT	885				

RESULT 6  
US-09-737-626A-10  
; Sequence 10, Application US/09737626A  
; Patent No. 6660911  
; GENERAL INFORMATION:  
; APPLICANT: Fincher, Karen  
; APPLICANT: Plasinski, Stanislaw  
; APPLICANT: Wilkinson, Jack  
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs  
; FILE REFERENCE: 38-21(51499)B  
; CURRENT APPLICATION NUMBER: US/09/737,626A  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 09/737,626  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 1271  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (1)..(1271)  
; OTHER INFORMATION: Y = t/u or c  
; OTHER INFORMATION: Act8 promoter polynucleotide sequence and intron  
US-09-737-626A-10

Query Match	5.2%;	Score	75.4;	DB	4;	Length	1271;
Best Local Similarity	51.6%;	Pred.	No.2.4e-10;				
Matches	214;	Conservative	3;	Mismatches	194;	Indels	4;
Qy	606	AAAAGCTGTTCTGCACAGCAGGTCA	CGTTA---TC	TTTTACCCTGGTGCGTAATGATTCGTG	662		
Db	471	AGAGGCTGTTCTGCACAGCCCATGTC	ACGTTATACTTTTTTCCG	TATGATCGAAATGATTCGTC	530		
Qy	663	TCGTGCAATTTTAATTATTTTTTGA	AAAGGCCGAAAAATAAGTTGT	TAAAGATATAAGATAAACCCG	722		
Db	531	TTTGYGAATTTAAATTATTTTCCA	AAATTTGAYGACTCTAAAGA	AAAAAATAAGTTTTTCA	590		
Qy	723	CCTATATAAATTTCATATATTTTCT	CCCCCGCTTTTGAAATTGCT	CGTTGCTCTCTCTCACTT	782		
Db	591	GATTAACCCGCTATATAAATAGTT	CAACACTCGGTTATTTCTTCT	CCCCCTCAAAGAT	650		

QY 962 CTGGATCTCGATCTGTTTCTCAATTCCTTGAGATCTGGAATCGTTAAATTT 1016  
|||||  
Db 1364 TTGGATCTCGATTACGATTTCTAAGTACCTTCAAAAGTTGTTCCGATTTGATT 1418  
|||||

## RESULT 8

US-09-737-626A-29  
; Sequence 29, Application US/09737626A  
; Patent No. 6660911  
; GENERAL INFORMATION:  
; APPLICANT: Fincher, Karen  
; APPLICANT: Flaksinski, Stanislaw  
; APPLICANT: Wilkinson, Jack  
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs  
; FILE REFERENCE: 38-21(51499)B  
; CURRENT APPLICATION NUMBER: US/09/737,626A  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 09/737,626  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 29  
; LENGTH: 1800  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (1)..(1800)  
; OTHER INFORMATION: y = t/u or c  
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 in  
; OTHER INFORMATION: tro  
US-09-737-626A-29

Query Match 5.1%; Score 73.8; DB 4; Length 1800;  
Best Local Similarity 51.3%; Pred. No. 7.7e-10;  
Matches 213; Conservative 3; Mismatches 195; Indels 4; Gaps 2;

QY 606 AAAAGCTGTCACAGCCAGGTCACGTTA--TCTTTACCTGTCGGTCAAAATGATTCGGT 662  
|||||  
Db 1004 AGAGGCTGTCACAGCCATGTCACGTTATCTTTTCCGTATGATCGAAATGATTCGTC 1063  
|||||  
QY 663 TCTGTCGATTTAAATTAATTTTTTTGAAAGCCGAAATAAAGTTGTAAGAGATAAACCG 722  
|||||  
Db 1064 TTGGYGAATTTAAATTTTCCAAATTCAGACTCTAAGAGAAAAAATAGTTTTTCA 1123  
|||||  
QY 723 CCTATATAATTCATATATTTCTCCCGCTTTGAAATGCTGTTGTCCTCCTCACTT 782  
|||||  
Db 1124 GATAAACCCGCTATATAAATAGTTCAACACTCGGTTTATTTCTCTCCCTCAAAGAT 1183  
|||||  
QY 783 TCATCAGCCGTTTGAATCTCCGGGACTTGACAGAGAGAACAGGAGAGACTAAGA 842  
|||||  
Db 1184 TGCTCTGCTCTTACGTTCTATCGGCCGTTGCAATTCCTCCGCGAATAAGAGAGAGAA 1243  
|||||  
QY 843 -GAGAAAGTAAGAGATAAATCCAGGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCA 901  
|||||  
Db 1244 GGAGAAAGTGAAGAGATCTTATCTGTCGTCGTTGTTCTTCTCTCGATCTCTCGA 1303  
|||||  
QY 902 TCTTCTTCTCCGCTCTTTCTTCCAGGTAATAGGAATCTTCTGATCTACTTTATTTG 961  
|||||  
Db 1304 TCTTCTGCTTTTGTCTTTCCGATTAAGGTAATTAATAACCTCCGATCTACTTGTCTTGTG 1363  
|||||  
QY 962 CTGGATCTCGATCTGTTTCTCAATTCCTTGAGATCTGGAATCGTTTAATTT 1016  
|||||  
Db 1364 TTGGATCTCGATTACGATTTCTAAGTTACCTTCAAAAGTTGTTTCCGATTTGATT 1418  
|||||

## RESULT 9

US-09-806-708B-22/c  
; Sequence 22, Application US/09806708B  
; Patent No. 6784342  
; GENERAL INFORMATION:  
; APPLICANT: The University of British Columbia  
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants

; FILE REFERENCE: 4810-58741  
; CURRENT APPLICATION NUMBER: US/09/806,708B  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/147,133  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22  
; LENGTH: 1141  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (1)..(1141)  
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters  
US-09-806-708B-22

Query Match 3.7%; Score 52.8; DB 4; Length 1141;  
Best Local Similarity 11.7%; Pred. No. 0.00031;  
Matches 112; Conservative 359; Mismatches 471; Indels 15; Gaps 4;

QY 150 TAGCAAAATTTAAACAACATTTTATGTATGCAAGAGTCACATATGTATATATTCATTC 209  
|||||  
Db 1002 TNNAMYYATTRWAAYAAAKWAGNNRMVYGAAGNKGWCAAMATGMBWADTAGKMC 943  
|||||  
QY 210 AGAATCGTTTTGACGAGTTCCGATCTAGTAGGACATTTATTTAATGTACATACTAATCG 269  
|||||  
Db 942 NNNNNWTTDVRMAMKAKNNNNNAYWTACYNRAATNNKMAHKWKTHGHSKRTTH 883  
|||||  
QY 270 TGAATAGTGATATGATGAACAATCTGTATCTTATGTATAAATATCATAAACACATCATG 329  
|||||  
Db 882 HTTCTCRRTKYNNNNNNARTVYVYHHAARRWNNAMWTRTNNNNNNNNNA-----CRNT 830  
|||||  
QY 330 AAAGACACTTCTTTCAGGGTCTGAATTAATATCATCAATCTTAATAGAAAACGAATT 389  
|||||  
Db 829 RTWAABWKSHSCNNNNNNNNNNNNNTWCHYTANABBCYRANNNNNAARMARTCNNYMH 770  
|||||  
QY 390 AAATACGTTGAATGTATGAAATCTAATTAATGAAACAAGCAACACGACGAGGACTAACT 449  
|||||  
Db 769 AAVTTHTDWCYKWTWNTWYHDMWTMTTBTITTRNMTTSTNNNNNNNNNNNNNNNNNM 710  
|||||  
QY 450 TGCCTGGATGTCGCGTTTAAAGTTAAACACATAAATAAAGCGAGTGTCAATCAACACGC 509  
|||||  
Db 709 WKAYAHATNNWGCWNNNTDARRTNTTVMRRRMTNTKTRWYSTTRRHHYTGATNNNNN 650  
|||||  
QY 510 GGATCGAGGTCACAGTCATGAGCCATCAAGCAAAAGAACTAATCCAGGGGTGAG 569  
|||||  
Db 649 NNNNNNNNNNSCCTCTRMWTRMTKGDGMTVRKVKWDRDTCYVDVWADSWVWVYAN 590  
|||||  
QY 570 ATGATTAAATAGTTTAAATAATTTAGTTAAACAGGAGAAAGCTGTCTGACAGCCAGGTCA 629  
|||||  
Db 589 WNRCDVYTRNTYCKSYAHSYWTWSNNAMRYRYRARNWSSMARWTRNNNNWMSGBVR 530  
|||||  
QY 630 CGTTATCTTTACCTGTGGTCGAAATGATTCGGTCTGTGATTTTAAATTTATTTTGTAA 689  
|||||  
Db 529 NERWAGTMMWRHNNNTDTRYVWKKWARBTITVYDSMCNAKSMWRGNMRAAKMWWAA 470  
|||||  
QY 690 AGCCGAAATAAAGTTGTAAGAGATAAAC-----CCGCTATATAAATTCATATATTTT 744  
|||||  
Db 469 NNDAGAMDHTVYMGNTMMMERAWKNNMNAWCCRRAVCCNNNNNRACVWHKHKMRWTK 410  
|||||  
QY 745 CTCTCCCGCTTTGAATTTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 804  
|||||  
Db 409 YMKWKAACNNNNKAMMYRVAMWVMSRDITNTDMMWMTSDWBHHWYTVDTYMRANNNNN 350  
|||||  
QY 805 GCGCATTTGACAGAGAAGAACCAAGGAAGAACTAAGAGAGA--AAGTAAGAGATAATCC 862  
|||||  
Db 349 NNWRBCKTTSWWMDHNMNTHCTYGNNTWGSAYBMAAAMSWAAGASBNVYVWVWRMTYM 290  
|||||  
QY 863 AGGATTCATTTCTCGGTTTGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 922  
|||||  
Db 289 GKTMTNNNNNNKAWYRTKTVAVWCNNRYYDVAWTKRNNYKYCYAYBMYBMVGMKHHW 230  
|||||

```

923  TTCAAGTAATAGAACTTTCTCGATCTAATTATTTGCTCGATCTCGATCTTGTTTC 982
          : : | : : : : : : : : : : : : : : : : : : : : :
229  BWRABRHSNNMMWVKCRNKTWVSWHYHAMRYB-KWABAVGCNNNNWKDRMAHHWCATN 171
          : : : : : : : : : : : : : : : : : : : : : : : :
983  TCAATTTCTTGGAGATCTGGAACTCGTTTAATTTGGAATCTGTGAACCTCAATAATCTT 1042
          : : : : : : : : : : : : : : : : : : : : : : : :
170  NNMGMWWTAYMHMHKKGAAWTNNKTABRDDHBAHVKTWYRWYDYWCMMCMMAKAKV 111
          : : : : : : : : : : : : : : : : : : : : : : : :
1043  TTGGTTTTACTAGAATCGATCTAAGTTGACCGATCAGTTAGCTCGATATATAGCTACC 1099
          : : : : : : : : : : : : : : : : : : : : : : : :
110  RTAMKHMWYITDRVYSANNTGVRWMMWCMWVYSMNNRWYRMRGRKTYWAMWYSMS 54
          : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-09-949-016-12699/c
; Sequence 12699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12699
; LENGTH: 19438
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12699

```

```

; Sequence 2, Application US/08451405A
; Patent No. 5736358
; GENERAL INFORMATION:
; APPLICANT: FASEL, NICOLAS JOSEPH
; APPLICANT: REYMOND, CHRISTOPHE DOMINIQUE
; TITLE OF INVENTION: DICTYOSTELID EXPRESSION VECTOR AND
; TITLE OF INVENTION: METHOD FOR EXPRESSING A DESIRED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE WEBB LAW FIRM
; STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE
; CITY: PITTSBURGH
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 15219-1818
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISK
; COMPUTER: Midwest Micro 486-50
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451.405A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,273
; FILING DATE: 15-JAN-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: UNKNOWN
; US-08-451-405A-2

Query March 3.4%; Score 48.6; DB 1; Length 731;
Best Local Similarity 49.0%; Pred. No. 0.0035;
Matches 129; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 168 TATTTTATGTGATCAGAGTCAGCATATGTATATTCATTCAGAAATCGTTTTGACGAGT 227
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 598 TTTTITTTTAAANAAAAAATAAATTTTGAATGNTTAAAGAAAAAATAATAAAAAAT 639
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 228 TCGGATGTAGTAGGACCATTTATTAATGTATACATACTAATCGTGAATAGTGATATGATGA 287
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 638 AATAATGTAGAAAAAGGTATTTTATTAAAAAGAAATTTATTACTACTATTAGGAAAA 579
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 288 AACATTGTATCTTATGTATATAATATCCATATAACACATCATGAAGACACATTTCTTTGAG 347
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 578 TATTTTATTTCTAATTTGATATATATAAATAAATAATAATAATGTTATTTGTTTGTAT 519
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 348 GGTCTGTAATTAATTTATGATCAATTCCTAATAGAAACGAATTAATAATTAGTTGAATTTGA 407
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 518 TAAAGGGTGTATGGTAAAAAANAANAANAANAATAAATAAATAAATAAATAAATAAATAA 459
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 408 TGAATCTAATTTGAACAAGCCAA 430
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 458 AAAATAATAATTAATAAATAA 436
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-949-016-17067/c
; Sequence 17067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17067
; LENGTH: 95255
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(95255)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17067

Query Match          3.3%; Score 48.2; DB 4; Length 95255;
Best Local Similarity 48.6%; Pred. No. 0.042;
Matches 160; Conservative 0; Mismatches 168; Indels 1; Gaps 1;

QY 90 TGGCTACACTCGATGAATGGTATCTCAACTCAACATTATATCTATACCAACATTAGT 149
Db 42648 TGTATATATATATATATATTTGTTATATATATATATATATATATATATATAT 42589

QY 150 TAGCAAAATTTAAACAACTATTTTATGTATGCAAGAGTCAGCATATGTATATATGATTC 209
Db 42588 TTGTATATATATATATATATTTCTGTATATAT-TATATTTGTATATATATATATATTTCTGTT 42530

QY 210 AGAATCGTTTTCACGAGTTCGAGTGPAGTGGCATTTATTTAATGTATACATCTAAATCG 269
Db 42529 ATATTATATTTTGTATATAATCTGTATATATATTTTGTATATAATATATAATATAA 42470

QY 270 TGAATAGTATATGATGAACATTTGTATCTTATTTGTATATAAATCCATAAACACATCATG 329
Db 42469 TTATATATATATATATATATTTATATATATATATATATATATATATATATATATAA 42410

QY 330 AAAGACACTTCTTTCAGGGTCTGAATTAATATATGATACAAATCTTAATAGAAAACGAAT 389
Db 42409 TATATAAATATATGATATATATTTATATATATATATATATATATATATATATATATACGATAT 42350

QY 390 AAATTACGTTGAATGTATGAAATCTAAT 418
Db 42349 ATATTATATATATATATGATATATATAT 42321

RESULT 13
US-09-257-770-1/c
; Sequence 1, Application US/09257770
; Patent No. 6306596
; GENERAL INFORMATION:
; APPLICANT: Lambowitz, Allen M.
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; APPLICANT: Mohr, Georg
; APPLICANT: Beall, Clifford J.
; TITLE OF INVENTION: Methods for Cleaving Single-Stranded and
; TITLE OF INVENTION: Double-Stranded DNA Substrates with Nucleotide
; TITLE OF INVENTION: Integrase
; FILE REFERENCE: 24671/04007
; CURRENT APPLICATION NUMBER: US/09/257,770
; CURRENT FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5203
; TYPE: DNA
; ORGANISM: S. cerevisiae
US-09-257-770-1

Query Match          3.3%; Score 47.8; DB 3; Length 5203;
Best Local Similarity 45.8%; Pred. No. 0.014;
Matches 160; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
```

```
QY 73 AATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAACTCAACATTATAC 132
Db 2378 AATATACCTCTATGTAATTTGTTTAAACATGATGATTTTCTACATCAATAGTTGAATTACAA 2319

QY 133 TTATACCAAAACATTAGTTAGCAAAATTTAAACAACATATTTTATGTATGCAAGAGTCAGC 192
Db 2318 ATACTACAAGGTTTATTAAAAATTAGCTTTAGCTGTAGGTAATATATATTTAATTGAATCG 2259

QY 193 ATATGTATAATTTGATTCAGAAATCGTTTGGACGAGTTCGGATGTAGTAGCCATTATTT 252
Db 2258 ATATATTCAAAAGGATCAGTTACTTTAGCTTCTGTATATATATATAAATATACCATGATTT 2199

QY 253 AATGTACATACTAATCGTGAATAGTATATGATGAACATTTGTATCTTTATTTGTATAAATA 312
Db 2198 TCAATTTTTCATATATTATCAAAAGTATTTCTTGGAAAATGGCAATTTAATTTATCAITTT 2139

QY 313 TCCATAAAACATCATGAAAGACACTTTCTTTCAGGGTCTGAATTAATTAATGATACAAATT 372
Db 2138 TCAATAATATATTAAATTTATAACCAAAATTTTTTAAATAGTTTACTTTAATGTTTTTAAATCTA 2079

QY 373 CTAATAGAAAACGAATTAATTAACGTTTGAATTTGTATGAAATCTAATTTGA 421
Db 2078 TATTTACTAGTAAAGTTAATACACATCAATGAATAATATAATACGTAATAGA 2030

RESULT 14
US-09-949-016-12199
; Sequence 12199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12199
; LENGTH: 109038
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(109038)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12199

Query Match          3.3%; Score 47.2; DB 4; Length 109038;
Best Local Similarity 53.2%; Pred. No. 0.084;
Matches 100; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 230 GGATCTAGTAGTAGCCATTATTTAATGTACATACATACTAATCGTGAATAGTATGATGAAA 289
Db 71807 GGATGAAGTTGGACCCCTTCTCAAATCACATACAACTCAAAATAGATAAATATACCTAAA 71866

QY 290 CATTGTATCTTATGTATATAATATCCATAAACACATCATGAAGACACTTTCTTTTCAGGG 349
Db 71867 TATAAAGGTAAAAATATAAAAAATCTTAGAAGAAAACAGGCATAAATCTTTGTGACATTTG 71926

QY 350 TCTGAATTAATTTATGATACAAATTTCTAATAGAAAACGAATTTAAATTCGTTGAATTTGATG 409
Db 71927 GGTTAAGCAATGATTTCTTAATAACACACCCGAAGAAAATAAGATAGATTGAATGTCATTT 71986

QY 410 AAATCTAA 417
Db 71987 GAAATTA 71994
```





*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 19:55:51 ; Search time 925 Seconds  
(without alignments)  
9584.350 Million cell updates/sec

Title: US-09-868-744B-1

Perfect score: 1444

Sequence: 1 attatgatctcaataacatt.....gtactggaatgtagtagtcc 1444

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 5700845 seqs, 3069779757 residues

Total number of hits satisfying chosen parameters: 11401690

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183.8	82.0	1220	19 US-10-688-745-13	Sequence 13, Appl
2	1174.8	81.4	1219	9 US-09-737-626A-9	Sequence 9, Appli
3	1174.8	81.4	1219	16 US-10-427-169-9	Sequence 9, Appli
4	1174.8	81.4	1219	16 US-10-427-180-9	Sequence 9, Appli
5	1174.8	81.4	1219	18 US-10-909-860-9	Sequence 9, Appli
6	1174.8	81.4	1219	19 US-10-920-869-9	Sequence 9, Appli
7	1174.8	81.3	1742	9 US-09-737-626A-30	Sequence 30, Appl
8	1174	81.3	1742	16 US-10-427-169-30	Sequence 30, Appl
9	1174	81.3	1742	16 US-10-427-180-30	Sequence 30, Appl
10	1174	81.3	1742	18 US-10-909-860-30	Sequence 30, Appl
11	1174	81.3	1742	19 US-10-920-869-30	Sequence 30, Appl

12	1114.6	77.2	12304	19 US-10-473-945-5	Sequence 5, Appli
13	1104.6	76.5	1259	9 US-09-887-384A-5	Sequence 5, Appli
14	1053	72.9	1202	9 US-09-887-384A-2	Sequence 2, Appli
15	1001.8	69.4	1342	9 US-09-887-384A-6	Sequence 6, Appli
16	950.2	65.8	1285	9 US-09-887-384A-3	Sequence 3, Appli
17	784.4	54.3	910	9 US-09-887-384A-4	Sequence 4, Appli
18	732.8	50.7	853	9 US-09-887-384A-1	Sequence 1, Appli
19	144	10.0	1468	17 US-10-425-114-14693	Sequence 14693, A
20	110	7.6	573	9 US-09-770-152-258	Sequence 258, App
21	94	6.5	94	9 US-09-887-384A-14	Sequence 14, Appl
22	84.4	5.8	97	9 US-09-887-384A-12	Sequence 12, Appl
23	79	5.5	79	9 US-09-887-384A-9	Sequence 9, Appli
24	77.4	5.4	79	9 US-09-887-384A-17	Sequence 17, Appli
25	76	5.3	2000	9 US-09-938-842A-3302	Sequence 3302, Ap
26	76	5.3	2000	11 US-09-938-842A-3302	Sequence 3302, Ap
27	75.4	5.2	1271	9 US-09-737-626A-10	Sequence 10, Appl
28	75.4	5.2	1271	16 US-10-427-169-10	Sequence 10, Appl
29	75.4	5.2	1271	16 US-10-427-180-10	Sequence 10, Appl
30	75.4	5.2	1271	18 US-10-909-860-10	Sequence 10, Appl
31	75.4	5.2	1271	19 US-10-920-869-10	Sequence 10, Appl
32	73.8	5.1	1800	9 US-09-737-626A-29	Sequence 29, Appl
33	73.8	5.1	1800	16 US-10-427-169-29	Sequence 29, Appl
34	73.8	5.1	1800	16 US-10-427-180-29	Sequence 29, Appl
35	73.8	5.1	1800	18 US-10-909-860-29	Sequence 29, Appl
36	73.8	5.1	1800	19 US-10-920-869-29	Sequence 29, Appl
37	65.4	4.5	67	9 US-09-887-384A-10	Sequence 10, Appl
38	65.4	4.5	74	9 US-09-887-384A-7	Sequence 8, Appli
39	63.4	4.4	77	9 US-09-887-384A-8	Sequence 2501, Ap
40	56.2	3.9	436	11 US-09-732-627A-2501	Sequence 7149, Ap
41	52.2	3.6	633	18 US-10-021-323-7149	Sequence 4034, Ap
42	51.6	3.6	580	18 US-10-767-795-4034	Sequence 3, Appli
43	50.8	3.5	1612	9 US-09-770-445-3	Sequence 1, Appli
44	50.6	3.5	3673778	16 US-10-312-841-1	Sequence 213, App
45	50.4	3.5	6070	15 US-10-311-455-213	

#### ALIGNMENTS

##### RESULT 1

; Sequence 13, Application US/10688745  
; Publication No. US20050086712A1  
; GENERAL INFORMATION:  
; APPLICANT: B.I. duPont de Nemours and Company Inc.  
; APPLICANT: Meyer, Knut  
; APPLICANT: Dhugga, Kanwarpal  
; TITLE OF INVENTION: Method to Produce para-Hydroxybenzoic Acid in the Stem Tissue of  
; FILE REFERENCE: a Plant  
; CURRENT APPLICATION NUMBER: US/10/688,745  
; CURRENT FILING DATE: 2003-10-17  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 13  
; LENGTH: 1220  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-688-745-13

Query Match	82.0%	Score 1183.8;	DB 19;	Length 1220;
Best Local Similarity	99.1%	Pred. No. 3.4e-271;		
Matches 1212;	Conservative	0;	Mismatches 7;	Indels 4; Gaps 2;
Qy	164	CAACTATTTTATGTCGACAGTCAGCATATGTAATTTGATTGATTCAGAACTGTTTCAC	223	
Db	1	CAACTATTTTATGTCGACAGTCAGCATATGTAATTTGATTGATTCAGAACTGTTTCAC	60	
Qy	224	GAGTTCGATGTCAGTATGAGCCATTTATTAATGTACATCTAATCGTGAATAGTCATG	283	
Db	61	GAGTTCGATGTCAGTATGAGCCATTTATTAATGTACATCTAATCGTGAATAGTCATG	120	
Qy	284	ATGAACACATTCGATCTTTATTTGTTATTAATAATATCCATAACACATCATGAAAGACATTTCTT	343	

Db 121 ATGAACATTTGATCTTTATTTGATATAATCATATAACACATCATGAAGAGACATTTCTTT 180  
QY 344 TCAGGCTCGAATTAATATGATACAAATCTTAATAGAAAACGAATTAATTAAGTTGAAT 403  
Db 181 TCAGGCTCTGAATTAATATGATACAAATCTTAATAGAAAACGAATTAATTAAGTTGAAT 240  
QY 404 TGTATGAATCTAAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTTGACT 463  
Db 241 TGTATGAATCTAAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTTGACT 300  
QY 464 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAAACACGCGATCGAGCAGGTC 523  
Db 301 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAAACACGCGATCGAGCAGGTC 360  
QY 524 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCAAGGGGTGAGATTAATTAAGTT 583  
Db 361 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCAAGGGGTGAGATTAATTAAGTT 420  
QY 584 TAAAAATTAAGTTAAACACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642  
Db 421 TAAAAATTAAGTTAAACACGAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 480  
QY 643 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGCCGAAAAATAA 702  
Db 481 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGCCGAAAAATAA 540  
QY 703 AGTTGAAGAGATAAACCCGCCATATAAATTCATATAATTTCTCCCGCTTTGAAATG 762  
Db 541 AGTTGAAGAGATAAACCCGCCATATAAATTCATATAATTTCTCCCGCTTTGAAATG 600  
QY 763 TCTCGTTGCTCCCTCACCTTCATCAGCGTTTGAATCTCCGGGACATTGACAGAGAG 822  
Db 601 TCTCGTTGCTCCCTCACCTTCATCAGCGTTTGAATCTCCGGGACATTGACAGAGAG 660  
QY 823 AACAGGAAGAAAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCCGTTT 882  
Db 661 AACAGGAAGAAAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCCGTTT 720  
QY 883 TGAATCTTCCCTCAATCTCTTCTTCTCCGCTCTTCTTCTTCCAGAGTAATAGGAAT 942  
Db 721 TGAATCTTCCCTCAATCTCA--TCTTCTTCCGCTCTTCTTCTTCCAGAGTAATAGGAAT 777  
QY 943 TCTGGATCTACTTTATTTCTGATCTCGATCTGTTTCTCAATTTCTCGATCTCG 1002  
Db 778 TCTGGATCTACTTTATTTCTGATCTCGATCTGTTTCTCAATTTCTCGATCTCG 837  
QY 1003 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAATCGAT 1062  
Db 838 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAATCGAT 897  
QY 1063 CTAAGTTGACCGATCAGTTAGTCGATTAATAGTACAGAAATTTGGCTTGACCTTGATGG 1122  
Db 898 CTAAGTTGACCGATCAGTTAGTCGATTAATAGTACAGAAATTTGGCTTGACCTTGATGG 957  
QY 1123 AGAGATCCATGTTTCATGTTTACCTGGGAATGATTTGTATATGTAATTCGAATCTGAAT 1182  
Db 958 AGAGATCCATGTTTCATGTTTACCTGGGAATGATTTGTATATGTAATTCGAATCTGAAT 1017  
QY 1183 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTTAGATTTGAATCTGAACACTGTTTAA 1242  
Db 1018 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTTAGATTTGAATCTGAACACTGTTTAA 1077  
QY 1243 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTTAGATTTGAATCTGAACACTGTTTAA 1302  
Db 1078 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTTAGATTTGAATCTGAACACTGTTTAA 1137  
QY 1303 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAATCTTTTGTG 1362  
Db 1138 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGATCTGATTTGAATCTTTTGTG 1197  
QY 1363 TGTGTCAGCTCATAAAAATGG 1385

Db 1198 TGTGTCAGCTCATAAACCATGG 1220  
RESULT 2  
US-09-737-626A-9  
; Sequence 9, Application US/09737626A  
; Patent No. US20020144304A1  
; GENERAL INFORMATION:  
; APPLICANT: Fincher, Karen  
; APPLICANT: Flasinski, Stanislaw  
; APPLICANT: Wilkenson, Jack  
; TITLE OF INVENTION: No. US20020144304A1el Plant Expression Constructs  
; FILE REFERENCE: 38-21(51499)B  
; CURRENT APPLICATION NUMBER: US/09/737,626A  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 09/737,626  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 1219  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (1)..(1219)  
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron  
US-09-737-626A-9

Query Match 81.4%; Score 1174.8; DB 9; Length 1219;  
Best Local Similarity 98.7%; Pred. No. 4.7e-269;  
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

QY 164 CAACTATTTTATGATGATCGAAGAGTCAGCATATGTAATTTGATTCAGATCGTTTGCAC 223  
Db 1 CAACTATTTTATGATGATCGAAGAGTCAGCATATGTAATTTGATTCAGATCGTTTGCAC 60  
QY 224 GAGTTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283  
Db 61 GAGTTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120  
QY 284 ATGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343  
Db 121 ATGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
QY 344 TCAGGCTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403  
Db 181 TCAGGCTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
QY 404 TGTATGAATCTAAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTTGACT 463  
Db 241 TGTATGAATCTAAATTTGAACAAGCCAAACACGACGAGGACTAACGTTGCTCGATTTGACT 300  
QY 464 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAAACACGCGATCGAGCAGGTC 523  
Db 301 CGGTTTAAAGTTAAACCACTAAAAACGAGCTGTCTATGTAAACACGCGATCGAGCAGGTC 360  
QY 524 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAAGTT 583  
Db 361 ACAGTCATGAAGCCATCAAGCAAGAAAGAACTAATCCAGGGGTGAGATTAATTAAGTT 420  
QY 584 TAAAAATTAAGTTAAACACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642  
Db 421 TAAAAATTAAGTTAAACACGAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 480  
QY 643 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGCCGAAAAATAA 702  
Db 481 TGTGTCGAAATGATTCGTCTGTCTGATTTTAAATTAATTTTGAAGCCGAAAAATAA 540  
QY 703 AGTTGAAGAGATAAACCCGCCATATAAATTCATATAATTTCTCCCGCTTTGAAATG 762  
Db 541 AGTTGAAGAGATAAACCCGCCATATAAATTCATATAATTTCTCCCGCTTTGAAATG 600



```
Db 1018 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA 1077
QY 1243 GTTAGATGAAGTTGTGTATAGATTCTTTCGAACCTTAGATTGTTAGTGTGTAGCTTTG 1302
Db 1078 GTTAGATGAAGTTGTGTATAGATTCTTTCGAACCTTAGATTGTTAGTGTGTAGCTTTG 1137
QY 1303 AACAGAAAGCTATTCTCTGATTCAATCAGGTTTATTTGACTGTATTGAACCTCTTTTGTG 1362
Db 1138 AACAGAAAGCTATTCTCTGATTCAATCAGGTTTATTTGACTGTATTGAACCTCTTTTGTG 1197
QY 1363 TGTTCGAGCTCATAAAAAATG 1384
Db 1198 TGTTCGAGCAGACTCACCATG 1219

RESULT 4
US-10-427-180-9
; Sequence 9, Application US/10427180
; Publication No. US20030199682A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiniski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. US20030199682A1e1 Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/427,180
; PRIOR FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: US/09/737,626A
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron
US-10-427-180-9
Query Match 81.4%; Score 1174.8; DB 16; Length 1219;
Best Local Similarity 98.7%; Pred.No. 4.7e-269;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;
QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGTATAATTGATTGAGAAATCGTTTGTAC 223
Db 1 CAACATATTTTATGATGCAAGAGTCAGCATATGTATAATTGATTGAGAAATCGTTTGTAC 60
QY 224 GAGTTCGGATGTAGTAGTACCCATTAATTAATGTACATATACTAATCGTGAATAGTATATG 283
Db 61 GAGTTCGGATGTAGTAGTACCCATTAATTAATGTACATATACTAATCGTGAATAGTATATG 120
QY 284 ATGAACATGTATCTTATGTATTAATATCCATAACATCATCAAGACATCTTCTT 343
Db 121 ATGAACATGTATCTTATGTATTAATATCCATAACATCATCAAGACATCTTCTT 180
QY 344 TCAGGCTCGAATTAATTAATGATACAAATCTTAATAGAAAACGAATTAATTAATTAATTAAT 403
Db 181 TCAGGCTCGAATTAATTAATGATACAAATCTTAATAGAAAACGAATTAATTAATTAATTAAT 240
QY 404 TGTATGAATCTTAATTAATGAACCCAAACCAACGACGAGCTAACGTTGCTGCTGATGACT 463
Db 241 TGTATGAATCTTAATTAATGAACCCAAACCAACGACGAGCTAACGTTGCTGCTGATGACT 300
QY 464 CGGTTTAAGTTAACCACTAAAAAAGCGAGCTGTCTATGTAACACGCGGATCCGAGCAGGTC 523
Db 301 CGGTTTAAGTTAACCACTAAAAAAGCGAGCTGTCTATGTAACACGCGGATCCGAGCAGGTC 360
QY 524 ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATTAATGTT 583
```

```
Db 361 ACAGTCATGAAGCCATCAAAAGCAAAAGAACTAATCCAAAGGGCTGAGATGATTAATTAATGTT 420
QY 584 TAAAAATTAGTTTAAACACGAGGGAAAAA-GCTGCTCAGACGCCAGGTCAGGTTATCTTTACC 642
Db 421 TAAAAATTAGTTTAAACACGAGGGAAAAAGGCTGCTCAGACGCCAGGTCAGGTTATCTTTACC 480
QY 643 TGTGTCGGAATGATTCGTGCTGCTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 702
Db 481 TGTGTCGGAATGATTCGTGCTGCTGATTTTAAATTAATTTTGAAGGCCGAAAAATAA 540
QY 703 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG 762
Db 541 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATG 600
QY 763 TCTCGTTGCTCTCCTCATTTCATCAGCGGTTTTGAATCTCCGCGACTTGACAGAGAG 822
Db 601 TCTCGTTGCTCTCCTCATTTCATCAGCGGTTTTGAATCTCCGCGACTTGACAGAGAG 660
QY 823 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCTCCGTTT 882
Db 661 AACAAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCTCCGTTT 720
QY 883 TGAATCTTCTCCTCAATCTCATCTTCTTCTTCCGCTCTTCTTCCAAAGGTAATAGAACTT 942
Db 721 TGAATCTTCTCCTCAATCTCA---TCCTTCTCCGCTCTTCTTCCAAAGGTAATAGAACTT 777
QY 943 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTCTGAGATCTGG 1002
Db 778 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTCTGAGATCTGG 837
QY 1003 AATTCTGTTTAAATTTGCGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAATCGAT 1062
Db 838 AATTCTGTTTAAATTTGCGATCTGTGAACCTCCACTAAATCTTTTGGTTTTACTAGAATCGAT 897
QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACCAAGAAATTTGGCTTTGACCTTGAATG 1122
Db 898 CTAAGTTGACCGATCAGTTAGCTCGATTAATAGCTACCAAGAAATTTGGCTTTGACCTTGAATG 957
QY 1123 AGAGATCCATGTTGATGTTACCTGGGAATGATTTGTATATGTAATGAAATCGAAATCGAACT 1182
Db 958 AGAGATCCATGTTGATGTTACCTGGGAATGATTTGTATATGTAATGAAATCGAAATCGAACT 1017
QY 1183 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTCAATCTGAACACTGTTTAA 1242
Db 1018 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTCAATCTGAACACTGTTTAA 1077
QY 1243 GTTAGATGAAGTTGTGTATAGATTCTTCGAAACCTTAGGATTTGATGTCGTAGCTTGG 1302
Db 1078 GTTAGATGAAGTTGTGTATAGATTCTTCGAAACCTTAGGATTTGATGTCGTAGCTTGG 1137
QY 1303 AACAGAAAGCTATTCTGATTCATCAATCAGGTTTTATTTGACTGTATTGAACCTCTTTTGTG 1362
Db 1138 AACAGAAAGCTATTCTGATTCATCAATCAGGTTTTATTTGACTGTATTGAACCTCTTTTGTG 1197
QY 1363 TGTTCGAGCTCATAAAAAATG 1384
Db 1198 TGTTCGAGCAGACTCACCATG 1219

RESULT 5
US-10-909-860-9
; Sequence 9, Application US/10909860
; Publication No. US20050005332A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiniski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: Novel Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/909,860
; CURRENT FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/737,626
```

```
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: promoter
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron
US-10-909-860-9

Query Match      81.4%; Score 1174.8; DB 18; Length 1219;
Best Local Similarity 98.7%; Pred. No. 4.7e-269;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGATATTAATGATTCAGAAATCGTTTGGAC 223
DB 1 CAACATATTTTATGATGCAAGAGTCAGCATATGATATTAATGATTCAGAAATCGTTTGGAC 60

QY 224 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
DB 61 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120

QY 284 ATGAACATCTGATCTTATTTGATTAATAATCAATCAATCAATCAATCAATCAATCAATCAAT 343
DB 121 ATGAACATCTGATCTTATTTGATTAATAATCAATCAATCAATCAATCAATCAATCAATCAAT 180

QY 344 TCAGGCTCGAATTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGAT 403
DB 181 TCAGGCTCGAATTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGAT 240

QY 404 TGTATGAATCTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGATGAT 463
DB 241 TGTATGAATCTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGATGAT 300

QY 464 CGGTTTAAGTTAAACCACTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 523
DB 301 CGGTTTAAGTTAAACCACTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 360

QY 524 ACAGTCATGAAGCCATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 583
DB 361 ACAGTCATGAAGCCATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420

QY 584 TAAATTAATGATTAACACGAGGGAAGAA-GCTGTCGACAGGAGGTCACGTTATCTTTTACC 642
DB 421 TAAATTAATGATTAACACGAGGGAAGAA-GCTGTCGACAGGAGGTCACGTTATCTTTTACC 480

QY 643 TGTGTCGAAATGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 702
DB 481 TGTGTCGAAATGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 540

QY 703 AGTTGTAAGAGATAAACCAGCCATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 762
DB 541 AGTTGTAAGAGATAAACCAGCCATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600

QY 763 TCTCGTTGTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 822
DB 601 TCTCGTTGTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 660

QY 823 AACAGGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCAATCAATCAATCAATCAATCAATCAAT 882
DB 661 AACAGGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCAATCAATCAATCAATCAATCAATCAAT 720

QY 883 TGAATCTTCTCCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
DB 721 TGAATCTTCTCCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 777

QY 943 TCTGATCTACTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCTTCTGAGATCTGG 1002
DB 943 TCTGATCTACTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCTTCTGAGATCTGG 1002

Db 778 TCTGATCTACTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCTTCTGAGATCTGG 837
QY 1003 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTCTTACTAGAAATCGAT 1062
DB 838 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTCTTACTAGAAATCGAT 897
QY 1063 CTAAGTTCACCGATCAGTTAGTCTCGATTTATAGTACCAAGAAATTTGGCTTGACCTTGATGG 1122
DB 898 CTAAGTTCACCGATCAGTTAGTCTCGATTTATAGTACCAAGAAATTTGGCTTGACCTTGATGG 957
QY 1123 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGATGATGATGATGATGATGAT 1182
DB 958 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGATGATGATGATGATGATGAT 1017
QY 1183 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTGTTAA 1242
DB 1018 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATGTTAGATTTGAATCTGAACACTGTGTTAA 1077
QY 1243 GTTAGATGAAGTTTGTGTATAGATTTCTTCGAAACCTTAGGATTTGTAGTGTCTGATGTTG 1302
DB 1078 GTTAGATGAAGTTTGTGTATAGATTTCTTCGAAACCTTTTAGGATTTGTAGTGTCTGATGTTG 1137
QY 1303 AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGCACTGTATTGAACCTCTTTTGTG 1362
DB 1138 AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGCACTGTATTGAACCTCTTTTGTG 1197
QY 1363 TGTTCGACGCTCATAAAAAATG 1384
DB 1198 TGTTCGACGACGACTCACCATG 1219

RESULT 6
US-10-920-869-9
; Sequence 9, Application US/10920869
; Publication No. US20050022261A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flaisinski, Stanislaw
; APPLICANT: Wilkison, Jack
; TITLE OF INVENTION: Novel Plant Expression Constructe
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/920,869
; CURRENT FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: US/09/737,626
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Act2 promoter polynucleotide sequence and intron
US-10-920-869-9

Query Match      81.4%; Score 1174.8; DB 19; Length 1219;
Best Local Similarity 98.7%; Pred. No. 4.7e-269;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGATATTAATGATTCAGAAATCGTTTGGAC 223
DB 1 CAACATATTTTATGATGCAAGAGTCAGCATATGATATTAATGATTCAGAAATCGTTTGGAC 60

QY 224 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
DB 61 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120

QY 284 ATGAACATCTGATCTTATTTGATTAATAATCAATCAATCAATCAATCAATCAATCAATCAAT 343
DB 121 ATGAACATCTGATCTTATTTGATTAATAATCAATCAATCAATCAATCAATCAATCAATCAAT 180

QY 344 TCAGGCTCGAATTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGAT 403
DB 181 TCAGGCTCGAATTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGAT 240

QY 404 TGTATGAATCTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGATGAT 463
DB 241 TGTATGAATCTAATTAATGATCAATCTAATAGAAACGAATTAATTAATGATGATGATGATGAT 300

QY 464 CGGTTTAAGTTAAACCACTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 523
DB 301 CGGTTTAAGTTAAACCACTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 360

QY 524 ACAGTCATGAAGCCATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 583
DB 361 ACAGTCATGAAGCCATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420

QY 584 TAAATTAATGATTAACACGAGGGAAGAA-GCTGTCGACAGGAGGTCACGTTATCTTTTACC 642
DB 421 TAAATTAATGATTAACACGAGGGAAGAA-GCTGTCGACAGGAGGTCACGTTATCTTTTACC 480

QY 643 TGTGTCGAAATGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 702
DB 481 TGTGTCGAAATGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 540

QY 703 AGTTGTAAGAGATAAACCAGCCATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 762
DB 541 AGTTGTAAGAGATAAACCAGCCATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600

QY 763 TCTCGTTGTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 822
DB 601 TCTCGTTGTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 660

QY 823 AACAGGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCAATCAATCAATCAATCAATCAATCAAT 882
DB 661 AACAGGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCAATCAATCAATCAATCAATCAATCAAT 720

QY 883 TGAATCTTCTCCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
DB 721 TGAATCTTCTCCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 777

QY 943 TCTGATCTACTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCTTCTGAGATCTGG 1002
DB 943 TCTGATCTACTTTATTTGCTGGATCTCGATCTTGTGTTTCTCAATTTCTTCTGAGATCTGG 1002
```









```
; CURRENT FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/737,626
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CaMV and Act2 polynucleotides + Act2 in
; OTHER INFORMATION: tro
US-10-909-860-30

Query Match      81.3%; Score 1174; DB 18; Length 1742;
Best Local Similarity 99.3%; Pred. No. 8.6e-269;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

QY 164 CAACATATTTTATGATGATCAAGAGTCAGCATATGATATTAATTCAGATCGTTTTCAC 223
DB 533 CAACATATTTTATGATGATCAAGAGTCAGCATATGATATTAATTCAGATCGTTTTCAC 592

QY 224 GAGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
DB 593 GAGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 652

QY 284 ATGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 343
DB 653 ATGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712

QY 344 TCAGGCTCGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 403
DB 713 TCAGGCTCGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 772

QY 404 TGTATGAATCTAAATGAAACGCAACACGAGGAGTAAACGCTGCTGCTGATGACT 463
DB 773 TGTATGAATCTAAATGAAACGCAACACGAGGAGTAAACGCTGCTGCTGATGACT 832

QY 464 CGTTTAAAGTTAAACCACTAAACCAACGAGCTGCTGATGATGATGATGATGATGATGAT 523
DB 833 CGTTTAAAGTTAAACCACTAAACCAACGAGCTGCTGATGATGATGATGATGATGATGAT 892

QY 524 ACAGTCATGAGCCATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 583
DB 893 ACAGTCATGAGCCATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 952

QY 584 TAAAAATAGTTAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 642
DB 953 TAAAAATAGTTAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1012

QY 643 TGTGTCGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
DB 1013 TGTGTCGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072

QY 703 AGTTCTAGAGATGAACCGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 762
DB 1073 AGTTCTAGAGATGAACCGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1132

QY 763 TCTCGTGTCTCTCTCACTTTCATCAGCGGCTTTTGAATCTCCGGGAGCTTGACAGAGAG 822
DB 1133 TCTCGTGTCTCTCTCACTTTCATCAGCGGCTTTTGAATCTCCGGGAGCTTGACAGAGAG 1192

QY 823 AACAGGAGAGAGACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 882
DB 1193 AACAGGAGAGAGACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1252

QY 883 TGAATCTCTCAATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
DB 1253 TGAATCTCTCAATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1309
```

```
RESULT 11
US-10-920-869-30
; Sequence 30, Application US/10920869
; Publication No. US2005002261A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiniski, Jack
; APPLICANT: Wilkenson, Stanislav
; TITLE OF INVENTION: Novel Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/920,869
; CURRENT FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: US/09/737,626
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CaMV and Act2 polynucleotides + Act2 in
; OTHER INFORMATION: tro
US-10-920-869-30

Query Match      81.3%; Score 1174; DB 19; Length 1742;
Best Local Similarity 99.3%; Pred. No. 8.6e-269;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

QY 164 CAACATATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 223
DB 533 CAACATATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 592

QY 224 GAGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
```

Db 593 GAGTTCCGATAGTAGTACCCATTATTTAATGTACATACATACTAATCGTGAATAGTGATG 652  
QY 284 ATGAACATTTGATCTTATTTGATATAATATCATATAACACATCATGAAGACACTTTCTT 343  
Db 653 ATGAACATTTGATCTTATTTGATATAATATCATATAACACATCATGAAGACACTTTCTT 712  
QY 344 TCAGGCTCTGAATTAATTTATGATACAAATTTCTAATAGAAAACGAATTTAAATTAAGTTGAAT 403  
Db 713 TCAGGCTCTGAATTAATTTATGATACAAATTTCTAATAGAAAACGAATTTAAATTAAGTTGAAT 772  
QY 404 TGTATGAATCTAATTTGAACAAGCCAAACACGACGAGGACTAAGCTTGCCTGGATGACT 463  
Db 773 TGTATGAATCTAATTTGAACAAGCCAAACACGACGAGGACTAAGCTTGCCTGGATGACT 832  
QY 464 CGGTTTAAGTTAAACCACTAAAAAAGCGAGCTGTATGTAAACACGCGGATCAGCAGGTC 523  
Db 833 CGGTTTAAGTTAAACCACTAAAAAAGCGAGCTGTATGTAAACACGCGGATCAGCAGGTC 892  
QY 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAGGGGTGAGATGATTAATTAAGTT 583  
Db 893 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAGGGGTGAGATGATTAATTAAGTT 952  
QY 584 TAAAAATTAAGTTAAACAGAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642  
Db 953 TAAAAATTAAGTTAAACAGAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 1012  
QY 643 TGTGTCGAAATGATTCGTGTCTGTGATTTTAAATTTATTTTGAAGGCCGAAAAATAA 702  
Db 1013 TGTGTCGAAATGATTCGTGTCTGTGATTTTAAATTTATTTTGAAGGCCGAAAAATAA 1072  
QY 703 AGTTGTAAGAGATAAACCAGCCCTATATATAATTCATATAATTTTCCCTCCCGCTTTGAAATG 762  
Db 1073 AGTTGTAAGAGATAAACCAGCCCTATATATAATTCATATAATTTTCCCTCCCGCTTTGAAATG 1132  
QY 763 TCTGTTGTCCTCCTCACTTTCATCAGCCGTTTGAATCTCCGGGACTGTGACAGAG 822  
Db 1133 TCTGTTGTCCTCCTCACTTTCATCAGCCGTTTGAATCTCCGGGACTGTGACAGAG 1192  
QY 823 AACAGGAGAGAGACTAAGAGAGAGTAAGAGATAATCCAGAGATTCATTTCCGTTT 882  
Db 1193 AACAGGAGAGAGACTAAGAGAGAGTAAGAGATAATCCAGAGATTCATTTCCGTTT 1252  
QY 883 TGAATCTTCTCAATCTCATCTTCTTCTTCCGCTCTTTCTTCCAGAGTAATAGGAACCTT 942  
Db 1253 TGAATCTTCTCAATCTCA--TCTTCTCCGCTCTTCTTCCAGAGTAATAGGAACCTT 1309  
QY 943 TCTGATCTACTTTATTTCTGATCTCGATCTTGTGTTTCTCAATTTCTTGAGATCTGG 1002  
Db 1310 TCTGATCTACTTTATTTCTGATCTCGATCTTGTGTTTCTCAATTTCTTGAGATCTGG 1369  
QY 1003 AATTGCTTTAATTTGGATCTGTGACCTCCACTTAATCTTTTGGTTTACTAGATTCGAT 1062  
Db 1370 AATTGCTTTAATTTGGATCTGTGACCTCCACTTAATCTTTTGGTTTACTAGATTCGAT 1429  
QY 1063 CTAAGTTGACCGATCAGTTAGTCTGATATAGTACACAGAAATTTGGCTTGAATCTGATGG 1122  
Db 1430 CTAAGTTGACCGATCAGTTAGTCTGATATAGTACACAGAAATTTGGCTTGAATCTGATGG 1489  
QY 1123 AGAGATCCATGTTGATTTAATCTCGGAAATGATTTGTATATGTAATGAAATCTGAAT 1182  
Db 1490 AGAGATCCATGTTGATTTAATCTCGGAAATGATTTGTATATGTAATGAAATCTGAAT 1549  
QY 1183 GTTGAAGTTAGATTTGAATCTGACCTGTCAATGTTAGATTCGAATCTGAATCTGAACT 1242  
Db 1550 GTTGAAGTTAGATTTGAATCTGACCTGTCAATGTTAGATTTGAATCTGAACTGTTAA 1609  
QY 1243 GTTAGATGAAGTTTGTGATAGATTTCTCGAAACCTTTAGGATTTGATGTCGTGATGG 1302  
Db 1610 GTTAGATGAAGTTTGTGATAGATTTCTCGAAACCTTTAGGATTTGATGTCGTGATGG 1669  
QY 1303 AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGTAATGAACTCTTTTGTG 1362  
Db 1670 AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGTAATGAACTCTTTTGTG 1729

QY 1363 TGTTTGCAGC 1372  
Db 1730 TGTTTGCAGC 1739

## RESULT 12

US-10-473-945-5  
; Sequence 5, Application US/10473945  
; Publication No. US20050081268A1  
; GENERAL INFORMATION:  
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE  
; APPLICANT: Kay, Steve A.  
; APPLICANT: Kuhlmann, Tina  
; APPLICANT: Lerner, Richard A.  
; TITLE OF INVENTION: BIOLUMINESCENT PLANTS AND METHODS OF MAKING SAME  
; FILE REFERENCE: SCRIP1340-1  
; CURRENT APPLICATION NUMBER: US/10/473,945  
; CURRENT FILING DATE: 2003-10-03  
; PRIOR APPLICATION NUMBER: PCT/US02/11116  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: US 60/282,094  
; PRIOR FILING DATE: 2001-04-06  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 12304  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: ACT-OM-LUC Vector  
US-10-473-945-5

Query Match 77.2%; Score 1114.6; DB 19; Length 12304;  
Best Local Similarity 98.0%; Pred. No. 2.9e-254;  
Matches 1150; Conservative 0; Mismatches 19; Indels 4; Gaps 2;

QY 224 GAGTTCCGATAGTAGTACCCATTATTTAATGTACATACTAATCGTGAATAGTGATG 283  
Db 1 GATCCCGATAGTAGTACCCATTATTTAATGTACATACTAATCGTGAATAGTGATG 60  
QY 284 ATGAACATTTGATCTTATTTGATATAATTCATATAACACATCATGAAAGACACTTTCTT 343  
Db 61 ATGAACATTTGATCTTATTTGATATAATTCATATAACACATCATGAAAGACACTTTCTT 120  
QY 344 TCAGGCTCTGAATTAATTTATGATACAAATCTTAATAGAAAACGAATTAATTAAGTTGAAT 403  
Db 121 TCAGGCTCTGAATTAATTTATGATACAAATCTTAATAGAAAACGAATTAATTAAGTTGAAT 180  
QY 404 TGTATGAATCTAATTTGAACCAAGCCAAACACGACGAGGACTAAACGTTGCTCGATTTGACT 463  
Db 181 TGTATGAATCTAATTTGAACCAAGCCAAACACGACGAGGACTAAACGTTGCTCGATTTGACT 240  
QY 464 CGGTTTAAAGTTAAACCACTAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCAGCAGGTC 523  
Db 241 CGGTTTAAAGTTAAACCACTAAAAAAGCGAGCTGTCTATGTAAACACGCGGATCAGCAGGTC 300  
QY 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAGGGGTGAGATGATTAATTAAGTT 583  
Db 301 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCCAAGGGGTGAGATGATTAATTAAGTT 360  
QY 584 TAAAAATTAAGTTAAACACGAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642  
Db 361 TAAAAATTAAGTTAAACACGAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 420  
QY 643 TGTGTCGAAATGATTCGTGTCTGTGATTTTAAATTTTGAAGGCCGAAAAATAA 702  
Db 421 TGTGTCGAAATGATTCGTGTCTGTGATTTTAAATTTTGAAGGCCGAAAAATAA 480  
QY 703 AGTTGTAAGAGATAAACCAGCCCTATATATAATTCATATAATTTTCCCTCCCGCTTTGAAATG 762  
Db 481 AGTTGTAAGAGATAAACCAGCCCTATATATAATTCATATAATTTTCCCTCCCGCTTTGAAATG 540

```
QY 763 TCTCGTTGTCCTCCTCACTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAAG 822
Db 541 TCTCGTTGTCCTCCTCACTTTCATCAGCCGTTTGAATCTCCGGGACCTTGACAGAGAAG 600
QY 823 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCAATCTCCGTTT 882
Db 601 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCAGGAGATTCAATCTCCGTTT 660
QY 883 TGAATCTTCTCAATCTCAATCTTCTTCTCCGCTCTTCTTCTTCCAAAGGTAATAGGAACCT 942
Db 661 TGAATCTTCTCAATCTCA ---TCTTCTTCCGCTCTTCTTCTTCCAAAGGTAATAGGAACCT 717
QY 943 TCTGGATCTACTTATTTCTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG 1002
Db 718 TCTGGATCTACTTATTTCTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG 777
QY 1003 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT 1062
Db 778 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT 837
QY 1063 CTAAGTTGACCGATCAGTTAGTCTGATATAGCTACCGAAATTTGGCTTGAACCTTGAAGG 1122
Db 838 CTAAGTTGACCGATCAGTTAGTCTGATATAGCTACCGAAATTTGGCTTGAACCTTGAAGG 897
QY 1123 AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATCGAAATCTGAACCT 1182
Db 898 AGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTATATGTGAATCGAAATCTGAACCT 957
QY 1183 GTTGAAGTTAGATTGAATCTGAACCTGCAACTGCAATGTTAGATTGAATCTGAACCTGTTTAA 1242
Db 958 GTTGAAGTTAGATTGAATCTGAACCTGCAACTGCAATGTTAGATTGAATCTGAACCTGTTTAA 1017
QY 1243 GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTGCTGAGTTG 1302
Db 1018 GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTGCTGAGTTG 1077
QY 1303 AACAGAAAGCTATTTCTGATTCATCAGAGGTTTATTTGATGATGAACTCTTTTGTG 1362
Db 1078 AACAGAAAGCTATTTCTGATTCATCAGAGGTTTATTTGATGATGAACTCTTTTGTG 1137
QY 1363 TGTTCAGCTCATAAAAATGCTCAGGCTGA 1395
Db 1138 TGTTCAGCTCAGGATCCATCGATAAGCTTTA 1170
```

## RESULT 13

```
US-09-887-384A-5
; Sequence 5, Application US/09887384A
; Patent No. US20020049992A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887, 384A
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1259
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Actin based
US-09-887-384A-5
```

```
Query Match 76.5%; Score 1104.6; DB 9; Length 1259;
Best Local Similarity 94.0%; Pred. No. 2.4e-252;
Matches 1198; Conservative 0; Mismatches 59; Indels 18; Gaps 4;
```

```
QY 164 CAACTATTTTATGATGCAAGAGTCAGCATATGATATTAATTCAGAAATCGTTTGTAC 223
Db 1 CAACTATTTTATGATGCAAGAGTTTCATCTTGTATATAATTCAGAAATCGTTTGTAC 60
```

```
QY 224 GAGTTCGGATGTAGTAGTACCCATTTATTAATGTACATCTAATCTGTGAATAGTGATAG 283
Db 61 GAGTTCGGATGTAGTAGTACCCATTTATTAATGTACATCTAATCTGTGAATAGTGATAG 120
QY 284 ATGAACAATTTGATCTTATTTGTATTAATATCCATAACAACATCATGAAGAAGACATTTCTT 343
Db 121 ATGAACAATTTGATCTTATTTGTATTAATATCCATAACAACATCATGAAGAAGACATTTCTT 180
QY 344 TCAGGGTCTGAAATTAATATATGATACAATTTCTAATAGAAAAACGAATTAATTAACGTTGAAT 403
Db 181 TCAGGGTCTGAAATTAATATATGATACAATTTCTAATAGAAAAACGAATTAATTAACGTTGAAT 240
QY 404 TGTATGAATCTAAATTTGAACCAACCAACGACGAGCACTAACGTTGCTCGATGATGACT 463
Db 241 TGTATGAATCTAAATTTGAACCAACCAACGACGAGCACTAACGTTGCTCGATGATGAA 300
QY 464 CGGTTTAAAGTTAAACCACTAAAAAAACGGAGCTGTCACTGTAAACACGCGGATCCAGCAGGTC 523
Db 301 AAGTTTAAAGATAAACCACTAAAAAAACGGAGCTGTCACTGTAAACACGCGGATCCAGCAGGTC 360
QY 524 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTTT 583
Db 361 ACAGTCATGAAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTTT 420
QY 584 TAAAAATTTAGTTAAACGAGGAAAAA - GCTGTCTGACAGCCAGGTCACGTTATCTTTACC 642
Db 421 TAAAAATTTAGTTAAACGAGGAAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTACC 480
QY 643 TGTGTGCAAAATGATTCGTGTCTGTGCTGATTTAAATTTATTTTTTGAAGGCCGAAATAA 702
Db 481 TGTGTGCAAAATGATTCGTGTCTGTGCTGATTTAAATTTATTTTTTGAAGGCCGAAATAA 540
QY 703 AGTGTGAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTG 762
Db 541 AGTGTGAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTG 600
QY 763 TCTCGTTGTCCTCCTCACTTTCATCAGCGGTTTGAATCTCCGGGACCTTGACAGAGAAG 822
Db 601 TCTCGTTGTCCTCCTCCTCACTTTCATCAGCGGTTTGAATCTCCGGGACCTTGACAGAGAAG 660
QY 823 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT 882
Db 661 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTT 706
QY 883 TGAATCTTCTCAATCTCAATCTTCTTCTTCCGCTCTTCTTCCCAAGGTAATAGGAACCT 942
Db 707 TGAATCTTCTCAATCTCA ---TCTTCTTCCGCTCTTCTTCCCAAGGTAATAGGAACCT 763
QY 943 TCTGGATCTACTTATTTTGTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG 1002
Db 764 TCTGGATCTACTTATTTTGTGCTGATCTCGATCTTGTGTTTCTCAATTTCTTGGATCTGG 823
QY 1003 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGAT 1062
Db 824 TTTTCGTTTAAATTTGGATGGAATTTAGATCACTAAATCTTTTGGTTTACTAGAAATCGAT 883
QY 1063 CTAAGTTGACCGATCAGTTAGTCTGATATAGCTACCGAAATTTGCTGTGACCTGTATGG 1122
Db 884 CTAAGTTGACCGATCAGTTAGTCTGATATAGCTACCGAAATTTGCTGTGACCTGTATGG 943
QY 1123 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCT 1182
Db 944 AGAGATCCATGTTTCATGTTTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACCT 1003
QY 1183 GTTGAAGTTAGATTGAATCTGAACCACTGTCAATGTTAGATTGAATCTGTAACCTGTTTAA 1242
Db 1004 GTTGAAGTTAGATTGAATCTGAACCACTGTCAATGTTAGATTGAATCTGTAACCTGTTTAA 1063
QY 1243 GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTGCTGAGTTG 1302
Db 1064 GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTGCTGAGTTG 1123
```

Qy	1303	AACAGAAGCTATTCTGATTCATCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1362
Db	1124	AACAGAAGCTATTCTGATTCATCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1183
Qy	1363	TGTTTGAGCTCATAAAAATGGCTGAGGCTGAGCATATTCAACCAATCGTGTGTGACAA	1422
Db	1184	TGTTTGAGCTCATAAACCATTGGCTGAGGCTGATGATTTCAACCAATCGTGTGTGACAA	1243
Qy	1423	TGTTACTGGAATGGT	1437
Db	1244	TGGTACTGGTATGGT	1258

RESULT 14

```

US-09-887-384A-2
; Sequence 2, Application US/09887384A
; Patent No. US2002004992A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887,384A
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Actin based
; OTHER INFORMATION: promoter
US-09-887-384A-2

```

Query Match	72.9%;	Score 1053;	DB 9;	Length 1202;
Best Local Similarity	94.0%;	Pred. No. 4.6e-240;		
Matches 1144;	Conservative 0;	Mismatches 55;	Indels 18;	Gaps 4;

Qy	703	AGTTGTAAGAGATAAACCCGCCCTATATAAAATTCATATATATTTCTCTCCCGCTTTGAATTG	762
Db	541	AGTTGTAAGAGATAAACCCGCCCTATATAAAATTCATATATATTTCTCTCTCCGCTTTGAATTG	600
Qy	763	TCTCGTTGTCTCTCTCATCTTTTCATCAGCCCGTTTTTGAATCTCCGCGCATTTGACAGAGAAG	822
Db	601	TCTCGTTGTCTCTCTCATCTTTTCATCAGCCCGTTTTTGAATCTCCGCGCATTTGACAGAGAAG	660
Qy	823	AACAAGGAAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATCTCCCGTTT	882
Db	661	AACAAGGAG-----AGAGAGAAAGTAAGAGATAATCCAGG-----TTCTCCGTTT	706
Qy	883	TGAATCTTCTCAATCTCATCTTCTCTCTCCGCTCTTTCTTCCAAAGTAAATAGGAACATT	942
Db	707	TGAATCTTCTCAATCTCA--TCTTCTCCGCTCTTTCTTCCAAAGTAAATAGGAACATT	763
Qy	943	TCTGGATCTACTTTATTTTGTCTGGATCTCGATCTTGTTTCTCAATTTCTTGGATCTGG	1002
Db	764	TCTGGATCTACTTTATTTTGTCTGGATCTCGATCTTGTTTCTCAATTTCTTGGATCTGG	823
Qy	1003	AAATTCGTTTAAATTTGGATCTGTGAACCTCCACATAAATCTTTTGGTTTTTACTAGAAATCGAT	1062
Db	824	TTTTTCGTTTAAATTTGGATGGAAATTAGATCACTAAATCTTTTGGTTTTTCTAGAAATCGAT	883
Qy	1063	CTAAGTTGACCGGATCAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTCGATGG	1122
Db	884	CTAAGTTGACCGGATCAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTCGATGG	943
Qy	1123	AGAGATCCATGTTCAATGTTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACT	1182
Db	944	AGAGATCCATGTTCAATGTTACCTGGGAAATGATTTGTATATGTGAATGAAATCTGAACT	1003
Qy	1183	GTTTCAAGTTAGATTGAATCTGAAACATGTCATGTTTGAATTTGAATCTGAAACATGTTTAA	1242
Db	1004	GTTTCAAGTTAGATTGAATCTGAAACATGTCATGTTTGAATTTGAATCTGAAACATGTTTAA	1063
Qy	1243	GTTTGAATGAAGTTTGTCTATAGATTCTTCCGAAACCTTAGGATTTGTAGTCTGTCGACGTTG	1302
Db	1064	GTTTGAATGAAGTTTGTCTATAGATTCTTCCGAAACCTTAGGATTTGTAGTCTGTCGACGTTG	1123
Qy	1303	AACAGAAAGCTATTTCTTGATTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1362
Db	1124	AACAGAAAGCTATTTCTTGATTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG	1183
Qy	1363	TGTTTGCAGCTCATAAA	1379
Db	1184	TGTTTGCAGCTCATAAA	1200

```

RESULT 15
US-09-887-384A-6
; Sequence 6, Application US/09887384A
; Patent No. US20020049992A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887,384A
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1342
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Actin based
; OTHER INFORMATION: promoter
US-09-887-384A-6

Query Match          69.4%;   Score 1001.8;   DB 9;   Length 1342;
Best Local Similarity 91.1%;   Pred. No. 7.3e-228;

```

Matches 1215; Conservative 0; Mismatches 62; Indels 56; Gaps 12;			
QY	158	TTTAAACAACACTATTTTATGTATGCAAGAGTCAGCATATGTATAATTTGATTGAGTACGATCGT	217
Db	12	TTGCATCAACTATTTTATGTATGCAAGAGTTTCATCTTGTATAATTTGATTGAGTACGATCGT	71
QY	218	TTTTCAGAGTTCCGATGTAGTAGTACCCATTTATTAATGTACATACTAATCGTGAATAGT	277
Db	72	TTTTCAGAGTTCCGATGTAGTAGTACCCATTTATTAATGTACATACTAATCGTGAATAGT	131
QY	278	GATATGATGAACAACATTTGATTTATTTGATATAATATCCATAAACAACATCATGAAGACAC	337
Db	132	GATATGATGAACAACATTTGATTTATTTGATATAATATCCATAAACAACATCATGAAGACAC	191
QY	338	TTTCTTTTCAGGGTCTGAATTAATTT-----ATGATACAATTTCTATAGAAAACGAATTA	390
Db	192	TTTCTTTTCAGGGTCTGAATTAATTTTGTAGAAATGATACAAATTTCTAATAGAAAACGAATTA	251
QY	391	AATTACGTTGAAATTTGATGAATCTAATTTGAACAGCCCAACCAACGACGAGGACTAACGTT	450
Db	252	AATTACGTTGAAATTTGATGAATCTAATTTGAACAGCCCAACCAACGACGAGGACTAACGTT	311
QY	451	GCCTGGATTGACTCGGTTTAAAGTTAAACCACTAAACCAACGAGCTGTATGTAAACACGCG	510
Db	312	GCCTGGAAATGAAGTTTAAAGTTAAACCACTAAACCAACGAGCTGTATGTAAACACGCG	371
QY	511	GATCAGCAGGTCACAGT-----CATGAAGCCATCAAAGCAAAAGAACTTAATCCAAAGGGG	565
Db	372	GATCAGCAGGTCACAGTCAATCCATGAAGCCATCAAAGCAAAAGAACTTAATCCAAAGGGG	431
QY	566	TCAGATGATTAATTTAGTTTAAAAATTTAGTTTAAACAGAGGGAAAA--GCTGTCAGACGCCA	624
Db	432	TCAGATGATTAATTTAGTTTAAAAATTTAGTATAACAGAGGGAAAAAGGCTGTCTGACAGCCA	491
QY	625	GGTCACGGTTA---TCTTTACCTGTGTGCAATGATTCGTGTCTGTGCGATTTTAAATATT	681
Db	492	GGTCACGGTTA---TCTTTACCTGTGTGCAATGATTCGTGTCTGTGCGATTTTAAATATT	551
QY	682	T-----TTTGAAGGCGGAAAAATAAAGTTGTAAAGATAAACCCGCGCTATATAA	731
Db	552	TCCAAAAATTGATTTTGAAGGCGGAAAAATAAAGTTGTAAAGATAAACCCGCGCTATATAA	611
QY	732	ATTCAAT-----ATATTTTCTCCCGCTTTGAAATTTGTCGTGTCCTCTCCTCACTT	782
Db	612	ATTCAATCAACACTCGATATTTTCTCTCCGCTTTGAAATTTGTCGTGTCCTCTCCTCACTT	671
QY	783	TCATCAGCGGTTTTCGATCTCCGCGACTTTCAGAGAGAACAGAGAGAGACTAAGA	842
Db	672	TCATCAGCGGTTTTCGATCTCCGCGACTTTCAGAGAGAACAGAGAGAGACTAAGA	731
QY	843	GAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCAT	902
Db	732	GAGAAAGTAAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATCTTCTCAATCTCA--	790
QY	903	CTTCTTTCTCCGCTCTTTCTTTCCAGGTAATAGGAACCTTTC--TGGAATCTACTTTATTG	961
Db	791	--TCTTTCTCCGCTCTTTCTTTCCAGGTAATAGGAACCTTTCTGGAATCTACTTTATTG	848
QY	962	CTGGATCTCGATCTTGTCTTCTCAATTTTCTTGAGATCTGGAATTCG--TTTAATTTGGA	1019
Db	849	CTGGATCTCGATCTTGTCTTCTCAATTTTCTTGAGATCTGGAATTTTAAATTTGGA	908
QY	1020	TCGTGGAACCTCCACTAAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGATCAG	1079
Db	909	TGGAATTTAGATCACTAATCTTTTGGTTTACTAGAAATCGATCTAAGTTGACCGATCAG	968
QY	1080	TTAGCTCGAATTAGCTACAGAAATTTGGCTTGAACCTTGATGGAGA---GATCCATGTTTC	1136
Db	969	TTAGCTCGAATTAGCTACAGAAATTTGGCTTGAACCTTGATGGAGATCCGATCCATGTTTC	1028
QY	1137	ATGTTACCTGGGAATGATTTGTATATGTGAATTTGAATCTGA--ACTGTTGAAGTTAGA	1194
Db	1029	ATGTTACCTGGGAATGATTTGTATATGTGAATTTGAATCTGACTACTGTTGAAGTTAGA	1088

Search completed: May 28, 2005, 00:05:31  
Job time : 932 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2005, 19:32:03 ; Search time 5118 Seconds  
(without alignments)  
10739.509 Million cell updates/sec

Title: US-09-868-744b-1  
Perfect score: 1444  
Sequence: 1 attatgatctcaataacatt.....gtactggaatggtaggatcc 1444

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	272.8	18.9	552	9	BX531590 Arabidops
2	202.4	14.0	327	1	AV831431
3	166	11.5	600	1	AV831862
4	163	11.3	546	1	AV525527
5	158	10.9	396	1	AV828177
6	158	10.9	562	1	AV830541
7	158	10.9	580	1	AV831261
8	158	10.9	632	5	BP561060
9	158	10.9	667	1	AV824265
10	158	10.9	671	5	BP560765
11	156	10.8	538	1	AV526249
12	153	10.6	399	1	AV829675
13	152	10.5	207	5	BP562156
14	151	10.5	297	1	AV525898
15	150	10.4	162	8	BH901724
16	147	10.2	774	5	BUC35286
17	146	10.1	463	7	N65062
18	144	10.0	438	7	N38049
19	144	10.0	597	7	T46735
20	143.8	10.0	382	1	AV545656
21	143	9.9	352	7	Z25952
22	143	9.9	393	7	T43759
23	142	9.8	411	7	R65270
24	139.4	9.7	253	7	T20889

## ALIGNMENTS

RESULT 1  
LOCUS BX531590/c 552 bp DNA linear GSS 04-APR-2004  
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-481G03-019880,  
Genomic survey sequence.  
ACCESSION BX531590  
VERSION BX531590.1 GI:31408720  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1  
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P., and Weissshaar, B.  
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana  
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)  
MEDLINE 22755829  
PUBMED 12874060  
REFERENCE 2  
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weissshaar, B.  
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics  
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)  
MEDLINE 23117147  
PUBMED 14756321  
REFERENCE 3  
AUTHORS Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and Weissshaar, B.  
TITLE High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines  
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)  
PUBMED 14682050  
REFERENCE 4  
AUTHORS Rosso, M.G., Li, Y., Strizhov, N. and Weissshaar, B.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAR-2004) Weissshaar B., Max-Planck-Institut fuer Zuechtungsforshung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
COMMENT This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At3g18780. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated

25 138.4 9.6 897 7 CF652111 34-102057  
26 136 9.4 454 1 AV525074 AV525074  
27 136 9.4 512 1 AV525969 AV525969  
28 136 9.4 518 1 AV525842 AV525842  
29 132.4 9.2 139 5 BP562323 BP562323  
30 132 9.1 370 7 T21821  
31 132 9.1 446 6 CB259737  
32 131 9.1 710 2 BE038458  
33 131 9.1 915 7 CF652259  
34 129.4 9.0 458 6 CB264030 44-L02036  
35 126 8.7 540 1 AV550855 AV550855  
36 125 8.7 449 7 CF651398 02-E92113-  
37 124 8.6 412 7 R29797 12402 Lambd  
38 123 8.5 426 1 AV554326 AV554326  
39 123 8.5 527 1 AV553948 AV553948  
40 123 8.5 546 1 AV518519 AV518519  
41 123 8.5 551 1 AV552230 AV552230  
42 120.6 8.4 251 6 CB258935 06-E01273  
43 119.8 8.3 270 1 AV519680 AV519680  
44 119 8.2 491 7 T76037 10815 Lambd  
45 119 8.2 502 7 N37237 18447 Lambd



'GABI'. Information on line availability can be found at:  
http://www.mpiz-koeln.mpg.de/GABI-Kat/.

```

FEATURES
  source
    1. .552
      Location/Qualifiers
        /organism="Arabidopsis thaliana"
        /mol_type="genomic DNA"
        /strain="Columbia 0"
        /db_xref="taxon:3702"
        /clone="CK-481G03-019880"
        /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
        /ecotype="Col-0"
        /note="PCR was performed on DNA from Arabidopsis thaliana
        plants (T1) which were transformed with the T-DNA from
        vector pAC161 (GenBank accession number: AJ537514). The
        lines contain one or more T-DNA insertions. The DNA
        fragment(s) resulting from the PCR were directly sequenced
        to determine the genomic sequence flanking the insertion.
        T-DNA derived sequences were removed."

ORIGIN
  Query Match      18.9%; Score 272.8; DB 9; Length 552;
  Best Local Similarity 93.9%; Pred. No. 3.3e-54;
  Matches 294; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 1126 GATCCATGTTTCATGTTACCTGGGAATGATTTGTATATGTAATCGAATCTGAACTGTT 1185
    |||||
Db 552 GATCCATGTTTCATGTTACCTGGGAATGATTTGTATATGTAATCGAATCTGAACTGTT 493

QY 1186 GAAGTTAGATTGAATCTGAACACTGTCGAATGTTAGATTGAATCTGAACACTGTTTAA-GT 1244
    |||||
Db 492 GAAGTTAGATTGAATCTGAACACTGTCGAATGTTAGATTGAATCTGAACACTGTTTAAAGGT 433

QY 1245 TAGATGAAGTTGTTGATAGATTCTCGAACTTCTAGGATTTGTAGTGTGCTACGTTGAA 1304
    |||||
Db 432 TAGATGAAGTTGTTGATAGATTCTCGAACTTCTAGGATTTGTAGTGTGCTACGTTGAA 373

QY 1305 CAGAAAGCTATTTCGATTCATTCAGGCTTTATTTGACTGTATTGAATCTCTTTTGTGTG 1364
    |||||
Db 372 CAGAAAGCTATTTCGATTCATTCAGGCTTTATTTGCTGTATTGAACTCTTTTGGGTG 313

QY 1365 TTTCGAGCTCATAAAAAATGGCTGAGCTGACGATATTCAACCAATCGTGTGTGACAATG 1424
    |||||
Db 312 TTTCGAGCTCATAAAAAATGGATGAGCTGATGATATTCAACCAATCGTGTGTGACAATG 253

QY 1425 GTACTGGAATGCT 1437
    |||||
Db 252 GTACCGGTATGCT 240

```

```

RESULT 2
AV831431
LOCUS
DEFINITION
  AV831431 RAFL9 Arabidopsis thaliana cdna clone RAFL09-89-B08 5',
  mRNA sequence.
ACCESSION
  AV831431
VERSION
  AV831431.1 GI:19873491
KEYWORDS
  EST.
SOURCE
  Arabidopsis thaliana (thale cress)
  ORGANISM
    Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
  1 (bases 1 to 327)
  Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
  Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
  Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
  and Shinozaki,K.
  Large scale analysis of Arabidopsis full-length cDNA (2002b)
  Unpublished (2002)
  Contact: Motoaki Seki
  Plant Functional Genomics Research Group
  RIKEN Genomic Sciences Center
  3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

REFERENCE
  AUTHORS
    TITLE
    JOURNAL
    COMMENT

```

Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: mseki@rtc.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
and XhoI was ligated to modified lambda PLC-1 vector (Carninci et  
al., submitted for publication) digested with BamHI and SalI. This  
clone is in a modified pBluescript vector. Please visit our web  
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further  
details.

```

FEATURES
  source
    1. .327
      Location/Qualifiers
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /db_xref="taxon:3702"
        /clone="RAFL09-89-B08"
        /dev_stages="plants at various developmental stages from
        germination to mature seeds"
        /lab_host="DH10B"
        /clone_lib="RAFL9"
        /note="Site 1: BamHI; Site 2: SalI; subjected to
        dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
        hr) treatments"

```

```

ORIGIN
  Query Match      14.0%; Score 202.4; DB 1; Length 327;
  Best Local Similarity 94.1%; Pred. No. 1.9e-37;
  Matches 222; Conservative 0; Mismatches 11; Indels 3; Gaps 1;

QY 716 AAACCCGGCTATATAATTCATATATTTTCCTCCCGCTTTGAATTTGTCCTGTTGCTCTC 775
    |||||
Db 3 AAACCCGGCTATATAATTCATATATTTTCCTCTCCGCTTTGAAATTTGTCCTGTTGCTCTC 62

QY 776 CTCACCTTCATCAGCCGTTTGAATCTCCGGCGACTTGACAGAGAACAGACAGAGAGAG 835
    |||||
Db 63 CTCACCTTCATCAGCCGTTTGAATCTCCGGCGACTTGACAGAGAACAGACAGAGAGAG 122

QY 836 ACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGAATCTTCCTCA 895
    |||||
Db 123 ACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGAATCTTCCTCA 182

QY 896 APTCATCTCTTCTTCCTCCGCTCTTTCTTTCCAGGTAATAGGAACCTTCTGATCT 951
    |||||
Db 183 ATCTCA---TCTTCTCCGCTCTTTCTTTCCAGGCTCATAAAAAATGGCTGAGGCT 235

```

```

RESULT 3
AV831862
LOCUS
DEFINITION
  AV831862 RAFL9 Arabidopsis thaliana cdna clone RAFL09-96-J07 5',
  mRNA sequence.
ACCESSION
  AV831862
VERSION
  AV831862.1 GI:19873922
KEYWORDS
  EST.
SOURCE
  Arabidopsis thaliana (thale cress)
  ORGANISM
    Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
  1 (bases 1 to 600)
  Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
  Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
  Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
  and Shinozaki,K.
  Large scale analysis of Arabidopsis full-length cDNA (2002b)
  Unpublished (2002)
  Contact: Motoaki Seki
  Plant Functional Genomics Research Group
  RIKEN Genomic Sciences Center
  3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
  Tel: 81-298-36-4359
  Fax: 81-298-36-9060
  Email: mseki@rtc.riken.go.jp

```

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

## FEATURES

source

Location/Qualifiers  
1..600  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAF09-96-J07"  
/dev\_stage="plants at various developmental stages from germination to mature seeds"  
/lab\_host="DH10B"  
/clone\_lib="RAF09"  
/note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) treatments"

## ORIGIN

Query Match 11.5%; Score 166; DB 1; Length 600;  
Best Local Similarity 93.4%; Pred. No. 1e-28;  
Matches 185; Conservative 0; Mismatches 10; Indels 3; Gaps 1;  
  
QY 754 TTTGAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTG 813  
|||||  
Db 1 TTTGAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTG 60  
  
QY 814 ACAGAGAACAAGGAAGAGACTAAGAGAAAGTAAGAGATAATCCAGAGATTTCAT 873  
|||||  
Db 61 ACAGAGAACAAGGAAGAGACTAAGAGAAAGTAAGAGATAATCCAGAGATTTCAT 120  
  
QY 874 TCTCCGTTTGAATCTTCTCAATCTCATCTTCTTCCGCTCTTCTTCCAGGTAA 933  
|||||  
Db 121 TCTCCGTTTGAATCTTCTCAATCTCA---TCCTTCTCCGCTCTTCTTCCAGGTCA 177  
  
QY 934 TAGGAACCTTCTGGATCT 951  
|||||  
Db 178 TAAAAATGGCTGAGGCT 195

## RESULT 4

AV525527  
LOCUS  
DEFINITION  
AV525527 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone APD2508R 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 546)  
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  
A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries  
DNA Res. 7 (3), 175-180 (2000)  
20363093  
PUBMED  
10907847  
Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: [asamizu@kazusa.or.jp](mailto:asamizu@kazusa.or.jp), URL: <http://www.kazusa.or.jp/en/plant/>.

## FEATURES

source

Location/Qualifiers  
1..546  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"

/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="APD2508R"  
/tissue\_type="aboveground organs"  
/dev\_stage="two to six-week old"  
/clone\_lib="Arabidopsis thaliana aboveground organs two to six-week old"  
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI"

## ORIGIN

Query Match 11.3%; Score 163; DB 1; Length 546;  
Best Local Similarity 93.3%; Pred. No. 5.3e-28;  
Matches 182; Conservative 0; Mismatches 10; Indels 3; Gaps 1;  
  
QY 757 GAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTGACA 816  
|||||  
Db 1 GAATTGTCGTGCTCCTCCTCATCTTCATCAGCGTTTGAATCTCCGGCACTTGACA 60  
  
QY 817 GAGAAGAACAGGAAGAGACTAAGAGAAAGTAAGAGATAATCCAGAGATTTCATCT 876  
|||||  
Db 61 GAGAAGAACAGGAAGAGACTAAGAGAAAGTAAGAGATAATCCAGAGATTTCATCT 120  
  
QY 877 CCGTTTGAATCTTCTCAATCTCATCTTCTTCCGCTCTTCTTCCAGGTAAATAG 936  
|||||  
Db 121 CCGTTTGAATCTTCTCAATCTCA---TCTTCTCCGCTCTTCTTCCAGGTCAATAA 177  
  
QY 937 GAACCTTCTGGATCT 951  
|||||  
Db 178 AAAATGGCTGAGGCT 192

## RESULT 5

AV828177  
LOCUS  
DEFINITION  
AV828177 RAF19 Arabidopsis thaliana cDNA clone RAF09-24-N21 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 396)  
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
Large scale analysis of Arabidopsis full-length cDNA (2002b)  
Unpublished (2002)  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: [mseki@rtc.riken.go.jp](mailto:mseki@rtc.riken.go.jp)  
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

## FEATURES

source

Location/Qualifiers  
1..396  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAF09-24-N21"  
/dev\_stage="plants at various developmental stages from

```

germination to mature seeds"
/lab_host="DH108"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: Sali; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Query Match      10.9%; Score 158; DB 1; Length 396;
Best Local Similarity 93.2%; Pred. No. 7.9e-27;
Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 762 GTCTGTTGTCCTCTCACTTTTCATCAGCGGTTTGAATCTCCGGCGACTTCACAGAGAA 821
Db 2 GTCTGTTGTCCTCTCACTTTTCATCAGCGGTTTGAATCTCCGGCGACTTCACAGAGAA 61
QY 822 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTT 881
Db 62 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTT 121
QY 882 TTGAATCTTCTCAATCTCATCTTCTTCTCCGCTCTTCTTCCAGGTAATAGGAAT 941
Db 122 TTGAATCTTCTCAATCTCA--TCTTCTCCGCTCTTCTTCCAGGCTCATAAAAT 178
QY 942 TTCTGGATCT 951
Db 179 GGCTGAGGCT 188

RESULT 6
AV830541
LOCUS AV830541 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-70-008 5',
DEFINITION mRNA sequence.
ACCESSION AV830541.1 GI:19872601
VERSION AV830541
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 562)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and Sali. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and Sali. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source
1..562
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone_lib="RAFL9-70-008"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH108"
/clone_lib="RAFL9"

/note="Site 1: BamHI; Site 2: Sali; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Query Match      10.9%; Score 158; DB 1; Length 396;
Best Local Similarity 93.2%; Pred. No. 7.9e-27;
Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 762 GTCTGTTGTCCTCTCACTTTTCATCAGCGGTTTGAATCTCCGGCGACTTCACAGAGAA 821
Db 2 GTCTGTTGTCCTCTCACTTTTCATCAGCGGTTTGAATCTCCGGCGACTTCACAGAGAA 61
QY 822 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTT 881
Db 62 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTT 121
QY 882 TTGAATCTTCTCAATCTCATCTTCTTCTCCGCTCTTCTTCCAGGTAATAGGAAT 941
Db 122 TTGAATCTTCTCAATCTCA--TCTTCTCCGCTCTTCTTCCAGGCTCATAAAAT 178
QY 942 TTCTGGATCT 951
Db 179 GGCTGAGGCT 188

RESULT 7
AV831261
LOCUS AV831261 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-87-G17 5',
DEFINITION mRNA sequence.
ACCESSION AV831261
VERSION AV831261.1 GI:19873321
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 580)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and Sali. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and Sali. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source
1..580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone_lib="RAFL09-87-G17"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH108"
/clone_lib="RAFL9"

/note="Site 1: BamHI; Site 2: Sali; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

```



```
|||||
62 GAACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCAATTCCTCGTT 121
|||||
882 TTGAATCTTCTCAATCTCATCTTCTTCTCGCTCTTCTTCCAAAGGTATAGGAAT 941
|||||
122 TTGAATCTTCTCAATCTCA---TCTTCTCGCTCTTCTTCTTCCAAAGCTCATAAAAAT 178
|||||
942 TTCTGATCT 951
|||||
179 GGCTGAGGCT 188

RESULT 10
LOCUS BP560765 671 bp mRNA linear EST 20-JUN-2004
DEFINITION BP560765 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-03-020 5', mRNA sequence.
ACCESSION BP560765
VERSION BP560765.1 GI:48976531
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T., Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M., Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
TITLE Functional annotation of a full-length Arabidopsis cDNA collection
JOURNAL Science 296 (5565), 141-145 (2002)
MEDLINE 21932900
PUBMED 11910074
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@tc.riken.go.jp
Please visit our web site (http://pfgweb.gsc.riken.go.jp/) for further details.
FEATURES
Location/Qualifiers
source
1..671
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-03-020"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated(1,2,5,10,24 hr)"

ORIGIN
Query Match 10.9%; Score 158; DB 5; Length 671;
Best Local Similarity 93.2%; Pred. No. 8.5e-27;
Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 762 GTCTGTTCTCTCTCACTTTTCATCAGCGGTTTTTGAATCTCGGCGACTTCACAGAGAA 821
|||||
2 GTCTGTTCTCTCTCACTTTTCATCAGCGGTTTTTGAATCTCGGCGACTTCACAGAGAA 61
|||||
822 GAACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTT 881
|||||
62 GAACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTT 121
|||||
882 TTGAATCTTCTCAATCTCATCTTCTTCTCGCTCTTCTTCCAAAGGTATAGGAAT 941
|||||
122 TTGAATCTTCTCAATCTCA---TCTTCTCGCTCTTCTTCTTCCAAAGCTCATAAAAAT 178
|||||
942 TTCTGATCT 951
|||||
```

```
|||||
179 GGCTGAGGCT 188

AV526249 538 bp mRNA linear EST 18-FEB-2004
AV526249 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone APZ08b09 5', mRNA sequence.
ACCESSION AV526249
VERSION AV526249.1 GI:8685777
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 538)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
source
1..538
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="APZ08b09R"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 10.8%; Score 156; DB 1; Length 538;
Best Local Similarity 93.1%; Pred. No. 2.5e-26;
Matches 175; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 764 CTCGTTGTCTCTCTCACTTTTCATCAGCGGTTTTTGAATCTCCGCGACTTCACAGAGAGA 823
|||||
1 CTCGTTGTCTCTCTCACTTTTCATCAGCGGTTTTTGAATCTCCGCGACTTCACAGAGAGA 60
|||||
824 ACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTTTT 883
|||||
61 ACAAGGAAGAACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCAATTCCTCGTTTT 120
|||||
884 GAATCTTCTCAATCTCATCTTCTTCTCGCTCTTCTTCCAAAGGTATAGGAATTT 943
|||||
121 GAATCTTCTCAATCTCA---TCCTTCTTCGCTCTTCTTCTTCCAAAGCTCATAAAAATGG 177
|||||
944 CTGGATCT 951
|||||
178 CTGAGGCT 185

RESULT 12
LOCUS AV829675 399 bp mRNA linear EST 01-APR-2002
DEFINITION AV829675 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-49-p05 5', mRNA sequence.
ACCESSION AV829675
VERSION AV829675.1 GI:19871735
```

EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 399)  
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,  
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.  
and Shinozaki, K.  
Large scale analysis of Arabidopsis full-length cDNA (2002b)  
Unpublished (2002)  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: msekic@r.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et  
al., submitted for publication) digested with BamHI and SalI. This  
clone is in a modified pBluescript vector. Please visit our web  
site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
details.

FEATURES  
source  
1..399  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAFL09-49-P05"  
/dev\_stage="plants at various developmental stages from  
germination to mature seeds"  
/lab\_host="DH10B"  
/clone\_lib="RAFL9"  
/note="Site 1: BamHI; Site 2: SalI; subjected to  
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
hr) treatments"

ORIGIN  
Query Match 10.6%; Score 153; DB 1; Length 399;  
Best Local Similarity 93.0%; Pred. No. 1.2e-25;  
Matches 172; Conservative 0; Mismatches 10; Indels 3; Gaps 1;  
QY 767 GTTGTCTCTCTCCTTCATCAGCGTTTGAATCTCGGCGACTTCACAGAGAACA 826  
DB 2 GTTGTCTCTCTCCTTCATCAGCGTTTGAATCTCGGCGACTTCACAGAGAACA 61  
QY 827 AGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTTGA 886  
DB 62 AGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTTGA 121  
QY 887 TCTTCTCAATCTCATCTTCTTCTTCGCTCTTCTTTCCTCAAGGTAATAGGAATCTTCTG 946  
DB 122 TCTTCTCAATCTCA--TCTTCTTCGCTCTTCTTTCCTCAAGCTCATATAAAATGGCTG 178  
QY 947 GATCT 951  
DB 179 AGGCT 183

RESULT 13  
BP562156  
LOCUS  
DEFINITION  
BP562156 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-56-H20 5',  
mRNA sequence.  
ACCESSION  
BP562156  
VERSION  
BP562156.1 GI:48977922  
KEYWORDS  
EST.  
Arabidopsis thaliana (thale cress)  
ORGANISM

EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 399)  
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,  
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.  
and Shinozaki, K.  
Large scale analysis of Arabidopsis full-length cDNA (2002b)  
Unpublished (2002)  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: msekic@r.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et  
al., submitted for publication) digested with BamHI and SalI. This  
clone is in a modified pBluescript vector. Please visit our web  
site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
details.

FEATURES  
source  
1..399  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAFL09-49-P05"  
/dev\_stage="plants at various developmental stages from  
germination to mature seeds"  
/lab\_host="DH10B"  
/clone\_lib="RAFL9"  
/note="Site 1: BamHI; Site 2: SalI; subjected to  
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
hr) treatments"

ORIGIN  
Query Match 10.6%; Score 153; DB 1; Length 399;  
Best Local Similarity 93.0%; Pred. No. 1.2e-25;  
Matches 172; Conservative 0; Mismatches 10; Indels 3; Gaps 1;  
QY 767 GTTGTCTCTCTCCTTCATCAGCGTTTGAATCTCGGCGACTTCACAGAGAACA 826  
DB 2 GTTGTCTCTCTCCTTCATCAGCGTTTGAATCTCGGCGACTTCACAGAGAACA 61  
QY 827 AGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTTGA 886  
DB 62 AGGAAGAAGACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTTGA 121  
QY 887 TCTTCTCAATCTCATCTTCTTCTTCGCTCTTCTTTCCTCAAGGTAATAGGAATCTTCTG 946  
DB 122 TCTTCTCAATCTCA--TCTTCTTCGCTCTTCTTTCCTCAAGCTCATATAAAATGGCTG 178  
QY 947 GATCT 951  
DB 179 AGGCT 183

RESULT 13  
BP562156  
LOCUS  
DEFINITION  
BP562156 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-56-H20 5',  
mRNA sequence.  
ACCESSION  
BP562156  
VERSION  
BP562156.1 GI:48977922  
KEYWORDS  
EST.  
Arabidopsis thaliana (thale cress)  
ORGANISM

EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 207)  
Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,  
Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M.,  
Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,  
Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.  
Functional annotation of a full-length Arabidopsis cDNA collection  
Science 296 (5565), 141-145 (2002)  
21932900  
11910074  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: msekic@r.riken.go.jp  
Please visit our web site (<http://pfweb.gsc.riken.go.jp/>) for  
further details.

FEATURES  
source  
1..207  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAFL09-56-H20"  
/dev\_stage="plants at various developmental stages from  
germination to mature seeds"  
/lab\_host="DH10B"  
/clone\_lib="RAFL9"  
/note="Site 1: BamHI; Site 2: SalI; subjected to  
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
hr) treatments"

ORIGIN  
Query Match 10.5%; Score 152; DB 5; Length 207;  
Best Local Similarity 98.2%; Pred. No. 1.9e-25;  
Matches 165; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 762 GTCTCGTTGTCTCTCCTTCATCAGCGTTTGAATCTCGGCGACTTCACAGAGAA 821  
DB 2 GTCTCGTTGTCTCTCCTTCATCAGCGTTTGAATCTCGGCGACTTCACAGAGAA 61  
QY 822 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTT 881  
DB 62 GAACAAGGAAGAACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTT 121  
QY 882 TTGAATCTTCTCAATCTCATCTTCTTCTTTCGCTCTTCTTTCCTCAAG 929  
DB 122 TTGAATCTTCTCAATCTCA--TCTTCTTTCGCTCTTCTTTCCTCAAG 166

RESULT 14  
AV525898  
LOCUS  
DEFINITION  
AV525898 Arabidopsis thaliana aboveground organs two to six-week  
old Arabidopsis thaliana cDNA clone APD3206R 5', mRNA sequence.  
ACCESSION  
AV525898  
VERSION  
AV525898.1 GI:8685426  
KEYWORDS  
EST.  
Arabidopsis thaliana (thale cress)  
ORGANISM

EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 297)  
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  
A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
of 12,028 non-redundant expressed sequence tags from normalized and  
size-selected cDNA libraries  
DNA Res. 7 (3), 175-180 (2000)  
20363093  
JOURNAL  
MEDLINE

PUBMED 10907847  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 252-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES  
 source  
 Location/Qualifiers  
 1..297  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="APD32b06R"  
 /tissue\_type="aboveground organs"  
 /dev\_stage="two to six-week old"  
 /clone\_lib="Arabidopsis thaliana aboveground organs two to six-week old"  
 /note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
 Query Match 10.5%; Score 151; DB 1; Length 297;  
 Best Local Similarity 92.9%; Pred. No. 3.5e-25;  
 Matches 170; Conservative 0; Mismatches 10; Indels 3; Gaps 1;  
 QY 769 TGTCTCTCTCCTTCATCAGCGGTTTTCGAATCTCCGGGACTTGACAGAGAAAGAACAG 828  
 Db 1 TGTCTCTCTCCTTCATCAGCGGTTTTCGAATCTCCGGGACTTGACAGAGAAAGAACAG 60  
 QY 829 GAAGAAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTGAATC 888  
 Db 61 GAAGAAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTTGAATC 120  
 QY 889 TTCCTCAATCTATCTTCTTCTTCCGCTCTTCTTCCAGGTAATAGGAATCTTCTGGA 948  
 Db 121 TTCCTCAATCTCA---TCTTCTTCCGCTCTTCTTCCAGGCTCATAAAAAATGGCTGAG 177  
 QY 949 TCT 951  
 Db 178 GCT 180

RESULT 15  
 BH901724/c  
 LOCUS 162 bp DNA linear GSS 04-SEP-2002  
 DEFINITION SALX\_085874.51.15.x Arabidopsis thaliana TDNA insertion lines.  
 Arabidopsis thaliana genomic clone SALX\_085874.51.15.x, genomic survey sequence.

ACCESSION BH901724.1 GI:22712605  
 VERSION BH901724  
 KEYWORDS GSS.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 162)  
 AUTHORS Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.  
 TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Joseph R. Ecker  
 Salk Institute Genomic Analysis Laboratory (SIGNAL)  
 The Salk Institute for Biological Studies  
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 Tel: 858 453 4100 x1752  
 Fax: 858 558 6379  
 Email: ecker@salk.edu  
 This is single pass sequence recovered from the left border of TDNA.  
 Class: TDNA tagged.

FEATURES  
 source  
 Location/Qualifiers  
 1..162  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /ecotype="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="SALK\_085874.51.15.x"  
 /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN  
 Query Match 10.4%; Score 150; DB 8; Length 162;  
 Best Local Similarity 99.4%; Pred. No. 5.5e-25;  
 Matches 161; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 575 TAATTAGTTTAAAAATTAGTTAAACACAGAGGAAAAA-GCTGTCTGACAGCCAGGTACCGTT 633  
 Db 162 TAATTAGTTTAAAAATTAGTTAAACACAGAGGAAAAAGGCTGTCTGACGCCAGGTACCGTT 103  
 QY 634 ATCTTTACCTGTGTGTCGAAATGATTTCGTCTGTCTGATTTTAAATTATTTTGTAAAGGC 693  
 Db 102 ATCTTTACCTGTGTGTCGAAATGATTTCGTCTGTCTGATTTTAAATTATTTTGTAAAGGC 43  
 QY 694 CGAAAAATAAGTTGTAAGAGATAAACCCGCCCTATATAATTC 735  
 Db 42 CGAAAAATAAGTTGTAAGAGATAAACCCGCCCTATATAATTC 1

Search completed: May 27, 2005, 23:45:17  
 Job time : 5123 secs